STIC-Biotech/ChemLib 117404

From: Rao, Manjunath N.
Sent: Monday, March 22, 2004 11:56 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/843,159

From: Manjunath N. Rao
Art Unit 1652, Room 3B-81
Mail Box in Room 3C-70
Phone: 272-0939

Date: 3-22-04

Please search the following as soon as possible for application with serial number **09/843,159**

SEQ ID NO: 1, and 2 against all **commercial nucleic acid databases** including **issued patents database** and **pending application database** and provide **a print of all results.**

SEQ ID NO:3 and 4 against all **commercial protein databases** including **issued patents database** and **pending application database** and provide **a print of all results.**

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao
Art Unit 1652, Room 3B81,
Mail Box in Room 3C70,
Remsen Building, USPTO
400, Dulany St.
Alexandria, VA.
Phone: 571-272-0939

RECEIVED
MAR 22 2004
(STIC)

Searcher: Arnold
Phone: 2-2532
Location:
Date Picked Up: 3/25/04
Date Completed: ~~3/24/04~~ 3/31/04
Searcher Prep/Review:
Clerical:
Online time:

TYPE OF SEARCH:
NA Sequences:
AA Sequences:
Structures:
Bibliographic:
Litigation:
Full text:
Patent Family:
Other:

VENDOR/COST (where applic.)
STN:
DIALOG:
Questel/Orbit:
DRLink:
Lexis/Nexis:
Sequence Sys.:
WWW/Internet:
Other (specify):

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:37:35 ; Search time 49.812 Seconds
(without alignments)
7854.389 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5760.5	89.1	1167	13	Q800E0	Q800e0 gallus gall
2	5074	78.5	1266	13	Q800D9	Q800d9 gallus gall
3	4285.5	66.3	976	11	Q8BX62	Q8bx62 mus musculu
4	4005	62.0	1181	5	Q9VBP3	Q9vbp3 drosophila
5	3997	61.8	1181	5	Q9XZ37	Q9xz37 drosophila
6	3062	47.4	600	11	Q8BXH7	Q8bxh7 mus musculu
7	904	14.0	1136	6	Q9N180	Q9n180 bos taurus
8	897	13.9	1887	4	Q7Z3G4	Q7z3g4 homo sapien
9	893	13.8	1726	11	Q8VC68	Q8vc68 mus musculu
10	891	13.8	1943	11	Q61307	Q61307 mus musculu
11	888.5	13.7	1762	11	O88521	O88521 rattus norv
12	885.5	13.7	2622	11	O70511	O70511 rattus norv
13	880.5	13.6	1863	4	Q7Z3L5	Q7z3l5 homo sapien
14	880	13.6	1856	4	Q99407	Q99407 homo sapien
15	878.5	13.6	1719	4	Q13768	Q13768 homo sapien
16	876	13.6	1848	11	Q61302	Q61302 mus musculu
17	872	13.5	1219	11	Q8C8R3	Q8c8r3 mus musculu
18	868.5	13.4	1549	5	Q9V4B1	Q9v4b1 drosophila
19	866.5	13.4	1549	5	Q24241	Q24241 drosophila
20	858	13.3	843	11	P97582	P97582 rattus norv
21	858	13.3	1809	5	Q17487	Q17487 caenorhabdi
22	858	13.3	1841	5	Q8MQG0	Q8mqg0 caenorhabdi
23	858	13.3	1867	5	Q17486	Q17486 caenorhabdi
24	858	13.3	2039	5	Q17489	Q17489 caenorhabdi
25	854	13.2	1815	5	Q17488	Q17488 caenorhabdi
26	854	13.2	6994	5	Q17343	Q17343 caenorhabdi
27	854	13.2	6994	5	Q17490	Q17490 caenorhabdi
28	851	13.2	1786	5	Q17344	Q17344 caenorhabdi
29	848.5	13.1	1159	5	Q9NCP8	Q9ncp8 drosophila
30	817	12.6	1009	5	Q8SWY2	Q8swy2 drosophila
31	724.5	11.2	1076	11	Q8BTI7	Q8bti7 mus musculu
32	714.5	11.1	919	4	Q8N8A2	Q8n8a2 homo sapien
33	705	10.9	1179	13	Q8AX92	Q8ax92 fugu rubrip
34	672	10.4	638	4	Q8IZ72	Q8iz72 homo sapien
35	662	10.2	1614	13	Q7T1G6	Q7t1g6 brachydanio
36	657	10.2	791	4	Q9UFA4	Q9ufa4 homo sapien
37	635.5	9.8	1619	5	Q9VMR4	Q9vmr4 drosophila
38	629.5	9.7	1088	4	Q13484	Q13484 homo sapien
39	627	9.7	1913	5	Q9GRV5	Q9grv5 caenorhabdi
40	590	9.1	1599	11	Q99NH0	Q99nh0 mus musculu
41	588.5	9.1	2486	4	O75179	O75179 homo sapien
42	563.5	8.7	934	16	O83807	O83807 treponema p
43	558.5	8.6	531	11	Q8C036	Q8c036 mus musculu
44	558.5	8.6	2542	4	Q8IWZ3	Q8iwz3 homo sapien
45	558.5	8.6	2617	4	Q8IWZ2	Q8iwz2 homo sapien

ALIGNMENTS

RESULT 1
Q800E0
ID Q800E0 PRELIMINARY; PRT; 1167 AA.
AC Q800E0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tankyrase 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA De Rycker M., Venkatesan R.N., Price C.M.;
RT "Vertebrate tankyrase domain structure and SAM-mediated
RT multimerization.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY142107; AAN41650.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 16.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 16.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 15.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 1167 AA; 126931 MW; 7842BF482B9229E9 CRC64;

Query Match 89.1%; Score 5760.5; DB 13; Length 1167;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1095; Conservative 33; Mismatches 36; Indels 3; Gaps 2;

```
Qy        75 MSGRRCAGGGAACASA--AAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRK 132
Db         1 MAARRCAGGAAALAEAPGCGSAVEP-ARELFEACRNGDVERVKRLVRPENVNSRDTAGRK  59

Qy       133 STPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPN 192
Db        60 SSPLHFAAGFGRKDVVEYLLQSGANVHARDDGGLIPLHNACSFGHAEVVNLLLRHGANPN 119

Qy       193 ARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKK 252
Db       120 ARDNWNYTPLHEAAIKGKTDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKK 179

Qy       253 DELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVH 312
Db       180 DELLESARSGNEEKMMSLLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVH 239

Qy       313 AKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLS 372
Db       240 AKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLS 299

Qy       373 YGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNF 432
Db       300 YGADPTLLNCHNKSTIDLAPTPQLKERLAYEFKGHSLLQAARESDVARIKKHLSLETVNF 359

Qy       433 KHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEV 492
Db       360 KHPQTHETALHCAAASPYPKRKQVCELLLRKGANINEKTKDFLTPLHVASEKAHNDVVEV 419

Qy       493 VVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQ 552
Db       420 VVKHEAKVNALDNLGQTSLHRAAHCGHLQTCRLLLSSGCDPSIVSLQGFTALQMGNESVQ 479

Qy       553 QLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNR 612
Db       480 QLLQEGIPLGNSDADRQLLEAAKAGDVDTVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNR 539

Qy       613 VSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHE 672
Db       540 VSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHE 599

Qy       673 AAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCL 732
Db       600 AAAKGKYEICKLLLQHGADPPKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCL 659

Qy       733 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 792
Db       660 ARVKKLCSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 719

Qy       793 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 852
Db       720 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 779

Qy       853 QTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAA 912
Db       780 QTPLDLVTADDVSALLTAAMPPSALPSCYKPQVISVSQTASSTADPLSSVPSSPSSLSAA 839

Qy       913 SSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ 972
Db       840 SSLDNLSGSFSELPSVVGTNSAEGATVLEKKEVSGVDFSINQFVRNLGLEHLIDIFEREQ 899

Qy       973 ITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDL 1032
Db       900 ITLDVLVEMGHKELKEIGINAYGHRHKIIKGVERLISGQQGLNPYLTLNTSSSGTLLIDL 959

Qy      1033 SPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSE 1092
Db       960 SSEDKEFQSVEEEMQSTVREHRDGGHAGGVFNRYNILKIQKVCNKKLWERYTHRRKEVSE 1019

Qy      1093 ENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGG 1152
Db       1020 ENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGG 1079

Qy      1153 TGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAE 1212
Db       1080 TGCPIHKDRSCYVCHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAE 1139

Qy      1213 YVIYRGEQAYPEYLITYQIMRPEGMVD 1239
Db       1140 YVIYRGEQAYPEYLITYQIVKPEATTE 1166
```

RESULT 2
Q800D9
ID Q800D9 PRELIMINARY; PRT; 1266 AA.
AC Q800D9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tankyrase 1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA De Rycker M., Venkatesan R.N., Price C.M.;
RT "Vertebrate tankyrase domain structure and SAM-mediated
RT multimerization.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY142108; AAN41651.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 17.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 18.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 15.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 1266 AA; 136211 MW; 14B8617EF930E6E0 CRC64;

Query Match 78.5%; Score 5074; DB 13; Length 1266;
Best Local Similarity 77.6%; Pred. No. 4.6e-313;
Matches 964; Conservative 109; Mismatches 120; Indels 50; Gaps 6;

```
Qy         5 RRGAAGGQG-----AQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLL  59
Db        49 RHSLAGPEGEAPPDAERPPAPECSEGAAPGP-PPGSGSSSSGSSASS--------------  93

Qy        60 LRLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAA-----RELFEACRNGDVERV  114
Db        94 ---------------SSSSSSTSSSVASSPAAESPEAAGPSGAFRELLEACRNGDVTRV  137

Qy       115 KRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACS  174
Db       138 KRLVDAGNVNAKDMAGRKSTPLHFAAGFGRKDVVEHLLQTGANVHARDDGGLIPLHNACS  197

Qy       175 FGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA  234
Db       198 FGHAEVVSLLLCQGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSA  257

Qy       235 LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAG  294
Db       258 LDLADPSAEAVLTGEYKKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAG  317

Qy       295 YNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTP  354
Db       318 YNRVRIVQLLLQHGADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTP  377
```

```
Qy        355 LHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAR 414
Db        378 LHEAASKNRVEVCSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAR 437

Qy        415 EADVTRIKKHLSLEMVNPKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEF 474
Db        438 EADLAKVKKTLALEIINFKQPQSHETALHCAVAAVHPKRKQVTELLLRKGANVNEKNKDF 497

Qy        475 LTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN 534
Db        498 MTPLHVAAEKAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLNYGSDPS 557

Qy        535 IISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCR 594
Db        558 IISLQGFTAAQIGNEAVQQILSESTPVRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCR 617

Qy        595 DIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVK 654
Db        618 DLEGRHSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVR 677

Qy        655 HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQ 714
Db        678 HGASVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQ 737

Qy        715 DLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ 774
Db        738 DLLRGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLE 797

Qy        775 HGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQL 834
Db        798 HGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQL 857

Qy        835 CALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGA 894
Db        858 CALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKLQAT-------V 910

Qy        895 TADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKK--EVPGVDFSI 952
Db        911 VSASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNTGDGAAGTERKEGEVSGLDMNI 970

Qy        953 TQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQ 1012
Db        971 TQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQ 1030

Qy       1013 GLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQ 1072
Db       1031 GTNPYLTFHCVSQGTILLDLAPDDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQ 1090

Qy       1073 KVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAG 1132
Db       1091 KVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGMFGAG 1150

Qy       1133 IYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHS 1192
Db       1151 IYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHA 1210

Qy       1193 PPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPE 1235
Db       1211 PPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIVKPE 1253
```

RESULT 3
Q8BX62
ID Q8BX62 PRELIMINARY; PRT; 976 AA.
AC Q8BX62;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE TRF1-interacting ankyrin-related ADP-ribose polymerase homolog.
GN TNKS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK048860; BAC33475.1; -.
DR PIR; PT0551; PT0551.
DR MGD; MGI:1341087; Tnks.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 15.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 13.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 12.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 976 AA; 106663 MW; 356B10CE9C33ECF1 CRC64;

Query Match 66.3%; Score 4285.5; DB 11; Length 976;
Best Local Similarity 82.0%; Pred. No. 3.5e-263;
Matches 798; Conservative 88; Mismatches 72; Indels 15; Gaps 3;

```
Qy        268 MALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACS 327
Db          1 MALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVPLHNACS 60

Qy        328 YGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSA 387
Db         61 YGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVNCHGKSA 120

Qy        388 IDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAA 447
Db        121 VDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETALHCAVA 180

Qy        448 SPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLG 507
Db        181 SLHPKRKQVAELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMNALDSLG 240

Qy        508 QTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEAD 567
Db        241 QTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPMRTSDVD 300

Qy        568 RQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVH 627
Db        301 YRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLHHGADVH 360

Qy        628 AKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQ 687
Db        361 AKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEICKLLLK 420

Qy        688 HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCR 747
Db        421 HGADPTKKNRDGNTPLDLVKEGDTDIQDLLRGDAALLDAAKKGCLARVQKLCTPENINCR 480

Qy        748 DTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 807
Db        481 DTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDIAALLIK 540

Qy        808 YNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSAL 867
Db        541 YNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATADDIRAL 600
```

```
Qy        868 LTAAMPPSALPSCYKPQ---VLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSE 924
Db        601 LIDAMPPEALPTCFKPQATVVSASVISPAST----------PSCLSAASSIDNLTGPLTD 650

Qy        925 LSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMG 982
Db        651 LAVGGASNAGDGAAGAERKEGEVAGLDMNIIQFLKSLGLEHLRDIFETEQITLDVLADMG 710

Qy        983 HKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSV 1042
Db        711 HEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEYQSV 770

Qy       1043 EEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERM 1102
Db        771 EEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHNERM 830

Qy       1103 LFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS 1162
Db        831 LFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRS 890

Qy       1163 CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAY 1222
Db        891 CYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAY 950

Qy       1223 PEYLITYQIMRPE 1235
Db        951 PEYLITYQIMKPE 963
```

```
RESULT 4
Q9VBP3
ID   Q9VBP3      PRELIMINARY;      PRT;  1181 AA.
AC   Q9VBP3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   BCDNA:LD22548 protein.
GN   BCDNA:LD22548 OR CG4719 OR CG17487.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003752; AAF56487.1; -.
DR   HSSP; Q00420; 1AWC.
DR   FlyBase; FBgn0027508; BcDNA:LD22548.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 17.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 14.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1181 AA;  127944 MW;  EB6F0F94E6A3D504 CRC64;
```

```
  Query Match             62.0%;  Score 4005;  DB 5;  Length 1181;
  Best Local Similarity   65.3%;  Pred. No. 3.1e-245;
  Matches  772;  Conservative  166;  Mismatches  189;  Indels   56;  Gaps   13;

Qy         91 AAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEY 150
Db         17 AVMANDP-LRELFEACKTGEIAKVKKLITPQTVNARDTAGRKSTPLHFAAGYGRREVVEF 75

Qy        151 LLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGK 210
Db         76 LLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYTPLHEAASKGK 135

Qy        211 IDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMAL 270
Db        136 VDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAARSGAEDRLLAL 195

Qy        271 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGH 330
Db        196 LTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLVPLHNACSYGH 255

Qy        331 YEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL 390
Db        256 FDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLLNCHSKSAIDA 315

Qy        391 APTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPY 450
Db        316 APTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDTPLHLAVVSPD 375

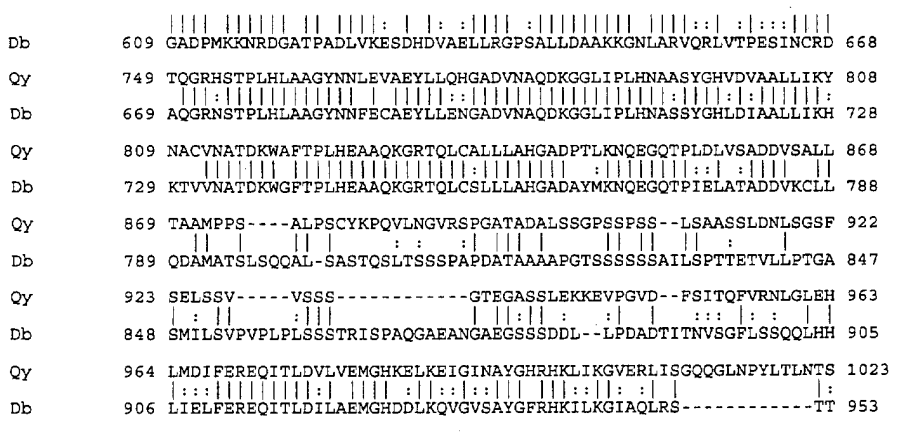
Qy        451 PKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTS 510
Db        376 GKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKVNALDSLGQTP 435

Qy        511 LHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQL 570
Db        436 LHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK-----NPPDSETHL 488

Qy        571 LEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 628
Db        489 LEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQFLLEHGAEVYA 548

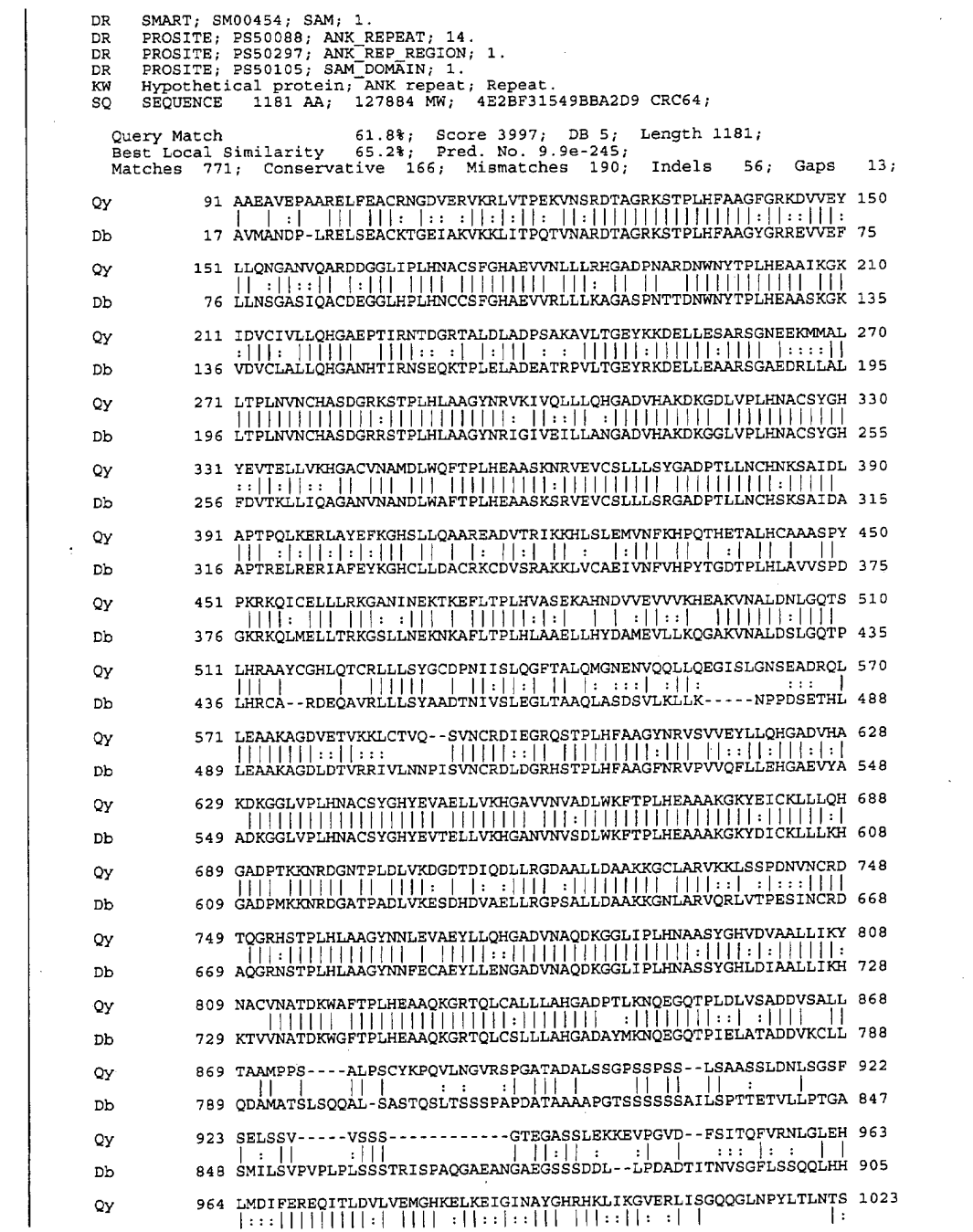
Qy        629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 688
Db        549 ADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGKYDICKLLLKH 608

Qy        689 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 748
```



RESULT 5
Q9XZ37
ID Q9XZ37 PRELIMINARY; PRT; 1181 AA.
AC Q9XZ37;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN BCDNA:LD22548 OR CG4719 OR CG17487.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF132196; AAD34784.1; -.
DR HSSP; Q00420; 1AWC.
DR FlyBase; FBgn0027508; BcDNA:LD22548.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 17.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 15.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 14.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein; ANK repeat; Repeat.
SQ SEQUENCE 1181 AA; 127884 MW; 4E2BF31549BBA2D9 CRC64;

Query Match 61.8%; Score 3997; DB 5; Length 1181;
Best Local Similarity 65.2%; Pred. No. 9.9e-245;
Matches 771; Conservative 166; Mismatches 190; Indels 56; Gaps 13;



```
Db        906 LIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS------------TT 953

Qy       1024 GSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK 1077
              | |       |:|:|| |||||| :|||||:|:|||| | ||| | ||||:::||| |:
Db        954 GIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYNIIRVQKVQNR 1013

Qy       1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
              |||||| |||:|::|||   :||||||||||:|||: :|||||||||||||||||||||
Db       1014 KLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGGMFGAGIYFAE 1073

Qy       1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
              :|||||||||||| ||| |||:|||:| |||| ||| ||||||:|||||||:|||||
Db       1074 HSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAMKMAHAPPGHH 1133

Qy       1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              || ||||  ||  ||||:|||||:||||||||::|:    |
Db       1134 SVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176


RESULT 6
Q8BXH7
ID   Q8BXH7      PRELIMINARY;      PRT;   600 AA.
AC   Q8BXH7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TANKYRASE-related protein (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK047094; BAC32960.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 14.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
FT   NON_TER     600    600
SQ   SEQUENCE   600 AA;  65307 MW;  2F667AD865F807F6 CRC64;

  Query Match             47.4%;  Score 3062;  DB 11;  Length 600;
  Best Local Similarity   97.2%;  Pred. No. 8e-186;
  Matches  583;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
              |||||||||||||||| ||||||:||||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAGAEAVEPSARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
              |||||||||||||||||||||||||||||||||||||||||||||||||:||| ||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLQHGAHPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:||:||
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGDYKEDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAEYNRVKIVQLLLHHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRIEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
              |||||||||||||||||| |||||:|||||||||||||||||||||||||||||||||||
Db        301 ADPTLLNCHNKSAIDLAPTAQLKERLSYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANTNEKTKEFLTPLHVASENAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
              |||||||||:|||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        421 KHEAKVNALDSLGQTSLHRAAHCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
              |||| |||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQEGASLGHSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600


RESULT 7
Q9N180
ID   Q9N180      PRELIMINARY;      PRT;  1136 AA.
AC   Q9N180;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ankyrin 1 (Fragment).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=breed Japanese Black; TISSUE=Bone marrow;
RA   Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;
RT   "Cloning of bovine erythrocyte ankyrin.";
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF222766; AAF61702.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER    1136   1136
SQ   SEQUENCE   1136 AA;  123182 MW;  07CD8B1568ACE8C7 CRC64;

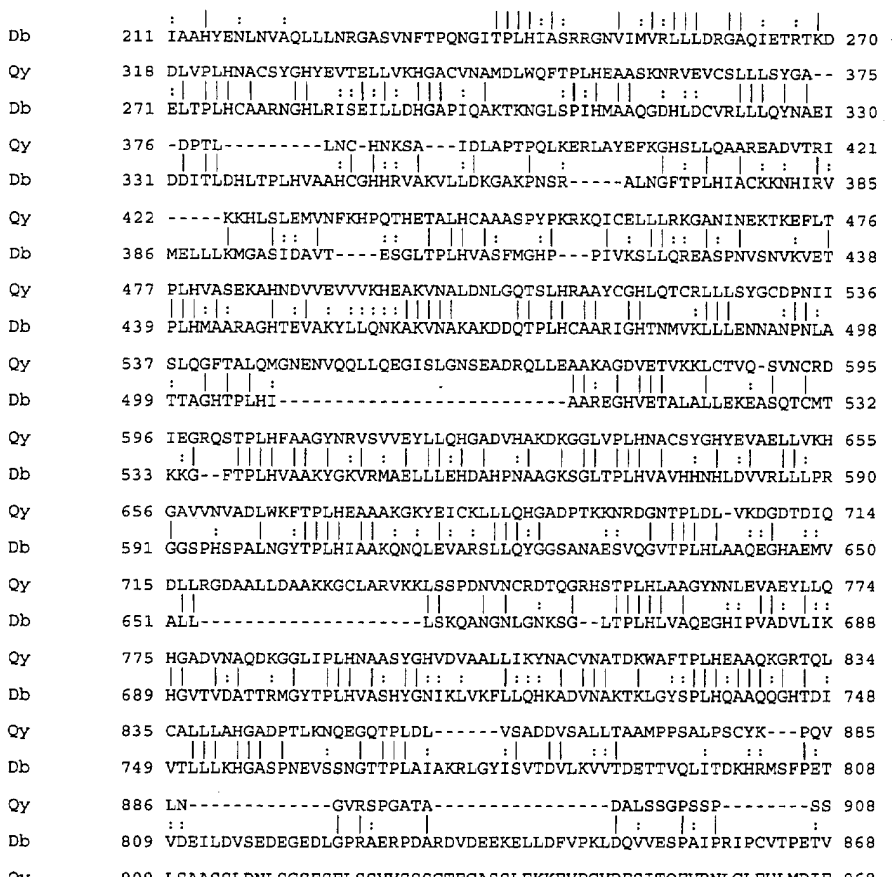
  Query Match             14.0%;  Score 904;  DB 6;  Length 1136;
  Best Local Similarity   26.9%;  Pred. No. 2.3e-48;
  Matches  296;  Conservative  162;  Mismatches  394;  Indels  250;  Gaps   29;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
              | : | |: |  |:  | :    |  : :| || ||  |: :||  |:  |||| |:
Db         51 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 109

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
              | ||: |   | |||  || :||: |      :|||  |  :|  :|   |: :| :
Db        110 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 169

Qy        225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
              :          || | |||         |||      ::|  :::|          |||:
Db        170 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 210

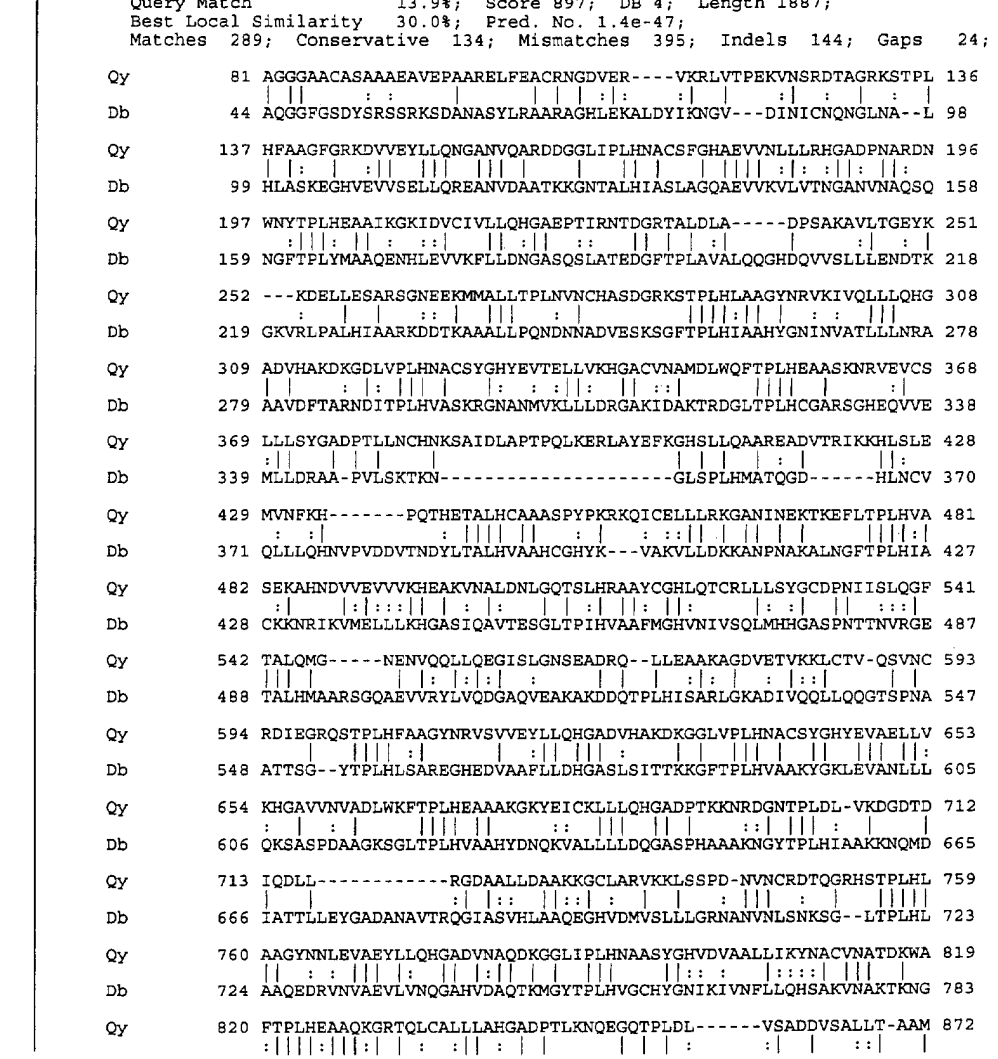
Qy        276 VNCHASDGRKS----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
```



RESULT 8
Q7Z3G4
ID Q7Z3G4 PRELIMINARY; PRT; 1887 AA.
AC Q7Z3G4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686P17114 (Fragment).
GN DKFZP686P17114.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human fetal kidney;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BX537917; CAD97900.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 1887 AA; 205502 MW; CAE428F76E8C5B5B CRC64;

Query Match 13.9%; Score 897; DB 4; Length 1887;
Best Local Similarity 30.0%; Pred. No. 1.4e-47;
Matches 289; Conservative 134; Mismatches 395; Indels 144; Gaps 24;



```
Db        784 YTPLHQAAQQGHTHIINVLLQNNASPNELTVNGNTALGIARRLGYISVVDTLKIVTEETM 843

Qy        873 PPSALPSCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVV 929

Db        844 TTTTVTEKHKMNVPETMNEVLD------------MSDDEVRKANAPEMLSDG-------- 883

Qy        930 SSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ-ITLDVLVEMGHKELKE 988

Db        884 --------------------------------EYISDVEEGEDAMTGDTDKYLGPQDLKE 911

Qy        989 IG 990

Db        912 LG 913


RESULT 9
Q8VC68
ID   Q8VC68      PRELIMINARY;      PRT;  1726 AA.
AC   Q8VC68;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   ANK3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC021657; AAH21657.1; -.
DR   MGD; MGI:88026; Ank3.
DR   GO; GO:0045202; C:synaptic junction; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   1726 AA;  188239 MW;  430148B67E00B1CA CRC64;

  Query Match           13.8%;  Score 893;  DB 11;  Length 1726;
  Best Local Similarity  28.3%;  Pred. No. 2.2e-47;
  Matches  337;  Conservative  167;  Mismatches  497;  Indels  188;  Gaps   38;

Qy         68 AVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRD 127

Db         10 AKPAHRKRKGKKSDANASYLRAARAGHLEKA----LDYIKNG----------VDVNICN 54

Qy        128 TAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRH 187

Db         55 QNGLNA--LHLASKEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTN 112

Qy        188 GADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSA 242

Db        113 GANVNAQSQNGFTPLYMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVV 172

Qy        243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299

Db        173 SLLLENDTKGKVRLPALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNIN 232

Qy        300 IVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAA 359

Db        233 VATLLLNRAAAVDFTARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGA 292

Qy        360 SKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVT 419

Db        293 RSGHEQVVEMLLDRSA-PILSKTKN----------------------GLSPLHMATQGD-- 328

Qy        420 RIKKHLSLEMVNFKH-------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTK 472

Db        329 ----HLNCVQLLLQHNVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKASPNAKAL 381

Qy        473 EFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCD 532

Db        382 NGFTPLHIACKKNRIRVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGAS 441

Qy        533 PNIISLQGFTALQMG-----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKL 585

Db        442 PNTTNVRGETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQL 501

Qy        586 CTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 644

Db        502 LQQGASPNAATTSG--YTPLHLAAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYG 559

Qy        645 HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLD 704

Db        560 KLEVASLLLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLH 619

Qy        705 L-VKDGDTDIQDLL------------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQ 750

Db        620 IAAKKNQMDIATSLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLSRNANVNLSNKS 679

Qy        751 GRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNA 810

Db        680 G--LTPLHLAAQEDRVNVAEVLVNQGAHVDAQTKMGYTPPHVGCHYGNIKIVNFLLQHSA 737

Qy        811 CVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS-------ADD 863

Db        738 KVNAKTKNGYTALHQAAQQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVDT 797

Qy        864 VSALLTAAMPPSALPSCYK---PQVLNGV--RSPGATADALSS------GP--------- 903

Db        798 LKVVTEEIMTTTTITEKHKMNVPETMNEVLDMSDDEGEDAITGDTDKYLGPQDLKELGDD 857

Qy        904 SSPSSLSAASSLDNLSGSFSELSSVVS-----SSGTEGASSLEKKEVPGVDFSITQFVRN 958

Db        858 SLPAEGYVGFSLGARSASLRSFSSDRSYTLNRSSYARDSMMIEELLVPSKEQHLT-FTRE 916

Qy        959 LGLEHLMDIFEREQITLDVL------VEMGHKELKEIGINAYG--------HRHKLI--- 1001

Db        917 FDSDSLRH-YSWAADTLDNVNLVSSPVHSGF--LVSFMVDARGGSMRGSRHHGMRIIIPP 973

Qy       1002 -------KGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS-VEEEMQSTVREH 1053

Db        974 RKCTAPTRITCRLVKRHKLANPPPMVEGEGLASRLVEMGPAGAQFLGPVIVEIP------ 1027

Qy       1054 RDGGHAGGIFNRYNILKIQKVCNKKLWE--RYTHRRKEVSE-----ENHNHANERMLFHG 1106

Db        1028 ----HFGSMRGKERELIVLRSENGETWKEHQFDSKNEDLAELLNGMDEELDSPEEL---G 1080

Qy       1107 SPFVNAIIHKGFDERHAYIGGMFGAGIYFAENS---SKSNQYVYGIGGG 1152

Db        1081 TKRICRIITKDFPQ------------YFAVVSRIKQESNQ--IGPEGG 1114


RESULT 10
Q61307
ID   Q61307      PRELIMINARY;      PRT;  1943 AA.
AC   Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; O08866; O08867;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ankyrin 3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin-3).
GN   ANK3.
```

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=95340633; PubMed=7615634;
RA   Peters L.L., John K.M., Lu F.M., Eicher E.M., Higgins A., Yialamas M.,
RA   Turtzo L.C., Otsuka A.J., Lux S.E.;
RT   "Ank3 (epithelial ankyrin), a widely distributed new member of the
RT   ankyrin gene family and the major ankyrin in kidney, is expressed in
RT   alternatively spliced forms, including forms that lack the repeat
RT   domain.";
RL   J. Cell Biol. 130:313-330(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RC   STRAIN=C3H/HEJ; TISSUE=Bone Marrow;
RX   MEDLINE=97213781; PubMed=9060470;
RA   Hoock T.C., Peters L.L., Lux S.E.;
RT   "Isoforms of ankyrin-3 that lack the NH2-terminal repeats associate
RT   with mouse macrophage lysosomes.";
RL   J. Cell Biol. 136:1059-1070(1997).
CC   -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
CC       SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
CC       BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
CC       AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
CC       DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
CC       THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC       THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC       ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=8;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=Q61307-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q61307-2; Sequence=VSP_050350, VSP_050353;
CC       Name=3;
CC         IsoId=Q61307-3; Sequence=VSP_050350, VSP_050353, VSP_050355;
CC       Name=4;
CC         IsoId=Q61307-4; Sequence=VSP_050354, VSP_050355;
CC       Name=5;
CC         IsoId=Q61307-5; Sequence=VSP_050352;
CC       Name=6;
CC         IsoId=Q61307-6; Sequence=VSP_050354;
CC       Name=7;
CC         IsoId=Q61307-7; Sequence=VSP_050349, VSP_050351, VSP_050355;
CC       Name=8;
CC         IsoId=Q61307-8; Sequence=VSP_050349, VSP_050351;
CC   -!- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC       EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC       LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC       SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC   -!- SIMILARITY: CONTAINS 24 ANK REPEATS.
DR   EMBL; L40631; AAB01603.1; -.
DR   EMBL; L40632; AAB01605.1; -.
DR   EMBL; U89275; AAB58381.1; -.
DR   EMBL; L40631; AAB01602.1; -.
DR   EMBL; L40632; AAB01604.1; -.
DR   EMBL; L40632; AAB01606.1; -.
DR   EMBL; L40632; AAB01607.1; -.
DR   EMBL; U89274; AAB58380.1; -.
DR   PIR; T42713; T42713.
DR   PIR; T42714; T42714.
DR   PIR; T42715; T42715.
DR   PIR; T42716; T42716.
DR   HSSP; P55273; 1BI8.
DR   MGD; MGI:88026; Ank3.
DR   GO; GO:0045202; C:synaptic junction; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Multigene family.
FT   DOMAIN        1    826       ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT   DOMAIN       23    810       24 X ANK MOTIF REPEATS.
FT   DOMAIN      856   1437       SPECTRIN BINDING DOMAIN.
FT   DOMAIN     1438   1943       REGULATORY DOMAIN (REGULATES BINDING OF
FT                                ANKYRIN TO SPECTRIN AND THE BAND 3
FT                                PROTEIN).
FT   REPEAT       23     55       ANK MOTIF 1.
FT   REPEAT       56     88       ANK MOTIF 2.
FT   REPEAT       89    121       ANK MOTIF 3.
FT   REPEAT      122    154       ANK MOTIF 4.
FT   REPEAT      155    183       ANK MOTIF 5.
FT   REPEAT      184    216       ANK MOTIF 6.
FT   REPEAT      217    249       ANK MOTIF 7.
FT   REPEAT      250    282       ANK MOTIF 8.
FT   REPEAT      283    315       ANK MOTIF 9.
FT   REPEAT      316    348       ANK MOTIF 10.
FT   REPEAT      349    381       ANK MOTIF 11.
FT   REPEAT      382    414       ANK MOTIF 12.
FT   REPEAT      415    447       ANK MOTIF 13.
FT   REPEAT      448    480       ANK MOTIF 14.
FT   REPEAT      481    513       ANK MOTIF 15.
FT   REPEAT      514    546       ANK MOTIF 16.
FT   REPEAT      547    579       ANK MOTIF 17.
FT   REPEAT      580    612       ANK MOTIF 18.
FT   REPEAT      613    645       ANK MOTIF 19.
FT   REPEAT      646    678       ANK MOTIF 20.
FT   REPEAT      679    711       ANK MOTIF 21.
FT   REPEAT      712    744       ANK MOTIF 22.
FT   REPEAT      745    777       ANK MOTIF 23.
FT   REPEAT      778    810       ANK MOTIF 24.
FT   VARSPLIC      1    805       Missing (in isoform 7 and isoform 8).
FT                                /FTId=VSP_050349.
FT   VARSPLIC      1    849       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_050350.
FT   VARSPLIC    833    853       Missing (in isoform 7 and isoform 8).
FT                                /FTId=VSP_050351.
FT   VARSPLIC    834    855       VRKASAPEKLSDGEYISDGEEG -> GDKCTWFKIPKVQEV
FT                                LVK (in isoform 5).
FT                                /FTId=VSP_050352.
FT   VARSPLIC    850    855       SDGEEG -> MALPHS (in isoform 2 and
FT                                isoform 3).
FT                                /FTId=VSP_050353.
FT   VARSPLIC    852    852       G -> GDKCTWFKIPKVQEVLVKS (in isoform 4
FT                                and isoform 6).
FT                                /FTId=VSP_050354.
FT   VARSPLIC   1569   1764       Missing (in isoform 3, isoform 4 and
FT                                isoform 7).
FT                                /FTId=VSP_050355.
SQ   SEQUENCE   1943 AA;  212013 MW;  B23C4B6F82904804 CRC64;

  Query Match             13.8%;  Score 891;  DB 11;  Length 1943;
  Best Local Similarity   30.6%;  Pred. No. 3.6e-47;
  Matches  287;  Conservative  127;  Mismatches  380;  Indels  144;  Gaps   24;

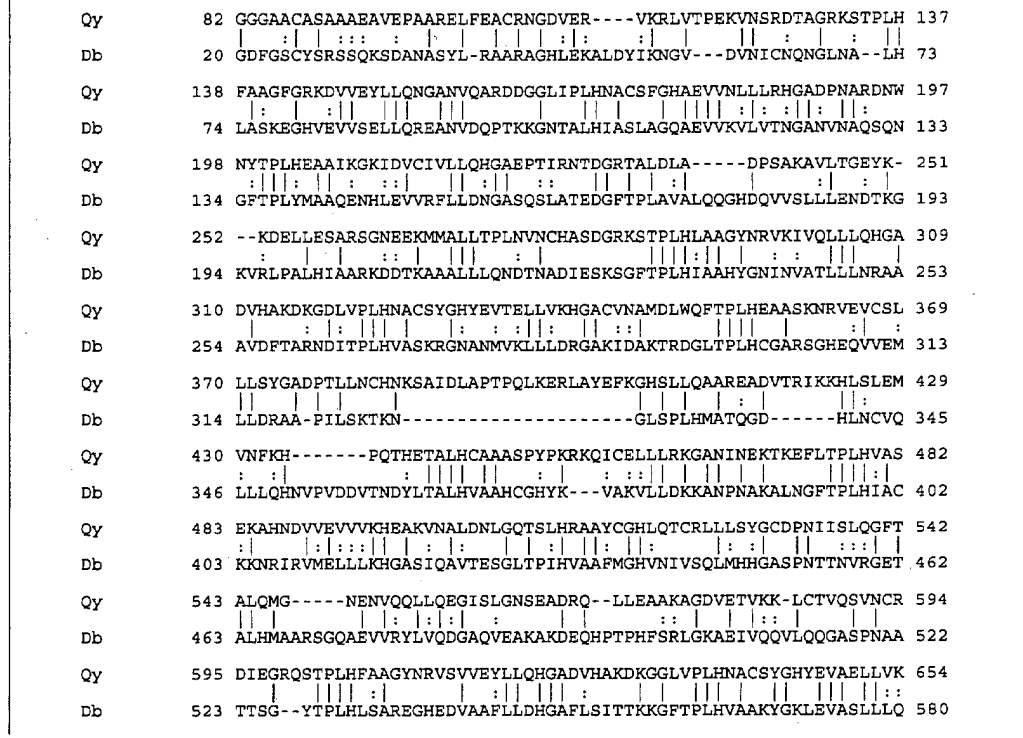
Qy        105 ACRNGDVER----VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
              | | | :|:    :| |    || : |  : || |:  |  :||  |||  ||| |
Db         31 AARAGHLEKALDYIKNGV---DVNICNQNGLNA--LHLASKEGHVEVVSELLQREANVDA 85

Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQH 220
                    |  || |   | |||| :|: :||: ||:    :|||: || :  ::|   || :
```



```
RESULT 11
O88521
ID   O88521      PRELIMINARY;      PRT;  1762 AA.
AC   O88521;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   190 kDa ankyrin isoform.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98395113; PubMed=9727010;
RA   Thevananther S., Kolli A.H., Devarajan P.;
RT   "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT   lung that associates with the plasma membrane and binds alpha-Na, K-
RT   ATPase.";
RL   J. Biol. Chem. 273:23952-23958(1998).
DR   EMBL; AF069525; AAC34809.1; -.
DR   HSSP; P55273; 1BI8.
DR   GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000169; SHprot_acsite.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 19.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1762 AA;  191899 MW;  9023280086A7BF4E CRC64;

  Query Match             13.7%;  Score 888.5;  DB 11;  Length 1762;
  Best Local Similarity   30.0%;  Pred. No. 4.4e-47;
  Matches  288;  Conservative  133;  Mismatches  396;  Indels  143;  Gaps   25;
```



```
Qy        655 HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDI 713
Db        581 KSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQMDI 640

Qy        714 QDLL-----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLA 760
Db        641 ATSLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLSRNANVNLSNKRGLN--PLHLG 698

Qy        761 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 820
Db        699 GQEDRVNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGY 758

Qy        821 TPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS-------ADDVSALLTAAMP 873
Db        759 TPLHQAAQQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVDTLKVVTEEIMT 818

Qy        874 PSALPSCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS 930
Db        819 TTTITEKHKMNVPETMNEVLD--MSDDEVGKG-SAPEKLSHGEYI-------------- 860

Qy        931 SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIG 990
Db        861 SDGEEG-------------------------------EDAITGDTDKSLGPQDLKELG 887
```

RESULT 12
O70511
ID O70511 PRELIMINARY; PRT; 2622 AA.
AC O70511;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).
GN ANK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98417633; PubMed=9744885;
RA Zhang X., Bennett V.;
RT "Restriction of 480/270-kD ankyrin G to axon proximal segments
RT requires multiple ankyrin G-specific domains.";
RL J. Cell Biol. 142:1571-1581(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Carpenter S.S., Zhang X.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 934-1220 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Skeletal muscle;
RX MEDLINE=98330444; PubMed=9664041;
RA Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RT "Ankyrin G is associated with the postsynaptic membrane and the
RT sarcoplasmic reticulum in the skeletal muscle fibre.";
RL J. Cell Sci. 111:2197-2207(1998).
DR EMBL; AF102552; AAC78143.1; -.
DR EMBL; AF065150; AAC18853.1; -.
DR HSSP; P55273; 1BI8.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
FT NON_TER 2622 2622
SQ SEQUENCE 2622 AA; 284456 MW; 67B34830D3AC884E CRC64;

Query Match 13.7%; Score 885.5; DB 11; Length 2622;
Best Local Similarity 30.5%; Pred. No. 1.3e-46;
Matches 283; Conservative 124; Mismatches 373; Indels 147; Gaps 24;

```
Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRK-STPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
Db         81 ASKEGHVEVVSELLQREA--NVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQ 138

Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
Db        139 NGFTPLYMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTK 198

Qy        224 PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDG 283
Db        199 GKVRLPALHIAARKDDTKAAALL--------LQNDTNADIESKMV-------VNRATESG 243

Qy        284 RKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGAC 343
Db        244 --FTSLHIAAHYGNINVATLLLNRAAAVDFTARNDITPLHVASKRGNANMVKLLLDRGAK 301

Qy        344 VNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYE 403
Db        302 IDAKTRDGLTPLHCGARSGHEQVVEMLLDRAA-PILSKTKN------------------- 341

Qy        404 FKGHSLLQAAREADVTRIKKHLSLEMVNFKH-------PQTHETALHCAAASPYPKRKQI 456
Db        342 --GLSPLHMATQGD------HLNCVQLLLQHNVPVDDVTNDYLTALHVAAHCGHYK---V 390

Qy        457 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 516
Db        391 AKVLLDKKANPNAKALNGFTPLHIACKKNRIRVMELLLKHGASIQAVTESGLTPIHVAAF 450

Qy        517 CGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQQLLQEGISLGNSEADRQ-- 569
Db        451 MGHVNIVSQLMHHGASPNTTNVRGETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTP 510

Qy        570 LLEAAKAGDVETVKKLCTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 628
Db        511 LHISARLGKADIVQQLLQQGASPNAATTSG--YTPLHLSAREGHEDVAAFLLDHGASLSI 568

Qy        629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 688
Db        569 TTKKGFTPLHVAAKYGKLEVASLLLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQ 628

Qy        689 GADPTKKNRDGNTPLDL-VKDGDTDIQDLL-----------RGDAALLDAAKKGCLARV 735
Db        629 GASPHAAAKNGYTPLHIAAKKNQMDIATSLLEYGADANPVTRQGIASVHLAAQEGHVDMV 688

Qy        736 KKLSSPD-NVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
Db        689 SLLLSRNANVNLSNKSG--LTPLHLAAQEDRVNVAEVLVNQGAHVDAQTKMGYTPLHVGC 746

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
Db        747 HYGNIKIVNFLLQHSAKVNAKTKNGYTPLHQAAQQGHTHIINVLLQNNASPNELTVNGNT 806

Qy        855 PLDLVS-------ADDVSALLTAAMPPSALPSCYK---PQVLNGVRSPGATADALSSGPS 904
Db        807 ALAIARRLGYISVVDTLKVVTEEIMTTTTITEKHKMNVPETMNEVLD------------M 854

Qy        905 SPSSLSAASSLDNLS-GSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEH 963
Db        855 SDDEVGKASAPEKLSDGEY-------ISDGEEG-------------------------- 880

Qy        964 LMDIFEREQITLDVLVEMGHKELKEIG 990
Db        881 ------EDAITGDTDKYLGPQDLKELG 901
```

```
RESULT 13
Q7Z3L5
ID   Q7Z3L5      PRELIMINARY;      PRT;  1863 AA.
AC   Q7Z3L5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein DKFZp686H0688.
GN   DKFZP686H0688.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Human retina;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX537758; CAD97827.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   1863 AA;  204736 MW;  1F4C998E0F0A03DF CRC64;

  Query Match             13.6%;  Score 880.5;  DB 4;  Length 1863;
  Best Local Similarity   29.3%;  Pred. No. 1.6e-46;
  Matches  308;  Conservative  145;  Mismatches  428;  Indels  171;  Gaps   26;

Qy       105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
                 | | |:::: |: |     :|: :  |  :  || ||  |   :|: ||  |::| :
Db        17 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 74

Qy       164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
                |   || |   | |||| :|:: ||: ||:    :|||: || :  |||   ||::||
Db        75 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 134

Qy       224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
                :    || |    ||       | |:|      | ::    |  |   :: |  |||
Db       135 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 192

Qy       276 VNCHASDGRKS-----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
                | | :| :               |||:|| |  | :  |||  || |    :  : |||
Db       193 -NDHNADIQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 251

Qy       325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
                |  |: : :||:  |  ::|       ||||  ||      :|  |||   || | | |
Db       252 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 310

Qy       385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
                                 | | |  | : |      |||     :  | |        :
Db       311 --------------------GLSPLHMAAQGDHVECVKHL----LQHKAPVDDVTLDYL 345

Qy       440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
                |||| ||    :    :: :||| | || | :      ||||:| :|    |:|::||: |
Db       346 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAS 402

Qy       500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN-----ENVQQL 554
                : |:   | | :| ||: |||     |||   |  |:: :::| ||| |        | |: |
Db       403 IQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCL 462

Qy       555 LQEGI-----------------SLGNSEADRQLLE-----------------AAKAGDV 579
                |: |                 || :|  : ||:                 :|: | |
Db       463 LRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTNGYTPLHISAREGQV 522

Qy       580 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
                :    |    : :   : :  |||| || |  : | : ||| |   :  | || |||
Db       523 DVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHV 581

Qy       640 ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDG 699
                |  | : :|| ||:: ||  :      :|||| || | : :|   || :||:     : |
Db       582 AAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNIVTKQG 641

Qy       700 NTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLH 758
                                                                        |||
Db       642 VTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----------------------TSLH 679

Qy       759 LAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
                |||  : : ||: | :|||| :|  | |  || |  ||:| :   |:|  | |||  |
Db       680 LAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYGNVKMVNFLLKQGANVNAKTKN 739

Qy       819 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLTAAM 872
                 :|||:|||:| | :  :|| ||| |       | | | :      :|  |   ::|  :
Db       740 GYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALAIAKRLGYISVVDTLKVVTEEV 799

Qy       873 PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSEL------- 925
               :      |    : :  |    : |          :|      | ::|   |
Db       800 TTTTTTITEK----HKLNVPETMTEVL--------DVSDEEGDDTMTGDGGEYLRPEDLK 847

Qy       926 ----SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RNLGLEHLMDIFEREQITLDVLV 979
                   |: ||   :| : |       |   |: |   |: | | |    : |:  :|  |
Db       848 ELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRSHTLSHAS--YLRDSAVMDDSV 905

Qy       980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
                :   ::  :   |  : ::|  | | |       |  |: :   ||      |  |: |
Db       906 VIPSHQVSTLAKEAERNSYRLSWGTENLD------NVALSSSPIHSGRASPCLERDNSSF 959

Qy      1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKI 1071
                  : | : : | | |    |:| |:|
Db       960 ------LVSFMVDARGGAMRG---CRHNGLRI 982


RESULT 14
Q99407
ID   Q99407      PRELIMINARY;      PRT;  1856 AA.
AC   Q99407;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ankyrin.
GN   ANK1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97382244; PubMed=9235914;
RA   Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;
RT   "Structure and organization of the human ankyrin-1 gene. Basis for
RT   complexity of pre-mRNA processing.";
RL   J. Biol. Chem. 272:19220-19228(1997).
DR   EMBL; U50133; AAB47805.1; -.
DR   EMBL; U50092; AAB47805.1; JOINED.
DR   EMBL; U50093; AAB47805.1; JOINED.
DR   EMBL; U50094; AAB47805.1; JOINED.
DR   EMBL; U50095; AAB47805.1; JOINED.
DR   EMBL; U50096; AAB47805.1; JOINED.
DR   EMBL; U50097; AAB47805.1; JOINED.
DR   EMBL; U50098; AAB47805.1; JOINED.
DR   EMBL; U50099; AAB47805.1; JOINED.
DR   EMBL; U50100; AAB47805.1; JOINED.
DR   EMBL; U50101; AAB47805.1; JOINED.
DR   EMBL; U50102; AAB47805.1; JOINED.
DR   EMBL; U50103; AAB47805.1; JOINED.
DR   EMBL; U50104; AAB47805.1; JOINED.
DR   EMBL; U50105; AAB47805.1; JOINED.
DR   EMBL; U50106; AAB47805.1; JOINED.
DR   EMBL; U50107; AAB47805.1; JOINED.
DR   EMBL; U50108; AAB47805.1; JOINED.
DR   EMBL; U50109; AAB47805.1; JOINED.
DR   EMBL; U50110; AAB47805.1; JOINED.
DR   EMBL; U50111; AAB47805.1; JOINED.
DR   EMBL; U50112; AAB47805.1; JOINED.
DR   EMBL; U50113; AAB47805.1; JOINED.
DR   EMBL; U50114; AAB47805.1; JOINED.
DR   EMBL; U50115; AAB47805.1; JOINED.
```

```
DR   EMBL; U50116; AAB47805.1; JOINED.
DR   EMBL; U50117; AAB47805.1; JOINED.
DR   EMBL; U50118; AAB47805.1; JOINED.
DR   EMBL; U50119; AAB47805.1; JOINED.
DR   EMBL; U50120; AAB47805.1; JOINED.
DR   EMBL; U50121; AAB47805.1; JOINED.
DR   EMBL; U50122; AAB47805.1; JOINED.
DR   EMBL; U50123; AAB47805.1; JOINED.
DR   EMBL; U50124; AAB47805.1; JOINED.
DR   EMBL; U50125; AAB47805.1; JOINED.
DR   EMBL; U50126; AAB47805.1; JOINED.
DR   EMBL; U50127; AAB47805.1; JOINED.
DR   EMBL; U50128; AAB47805.1; JOINED.
DR   EMBL; U50129; AAB47805.1; JOINED.
DR   EMBL; U50130; AAB47805.1; JOINED.
DR   EMBL; U50131; AAB47805.1; JOINED.
DR   EMBL; U50132; AAB47805.1; JOINED.
DR   HSSP; Q00420; 1AWC.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1856 AA;  203466 MW;  4A4E3A794DD6401F CRC64;

  Query Match           13.6%;  Score 880;  DB 4;  Length 1856;
  Best Local Similarity  29.9%;  Pred. No. 1.7e-46;
  Matches  254;  Conservative  131;  Mismatches  284;  Indels  180;  Gaps   22;

Qy         98 AARELFEACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGA 156
              ||      | |:|:::: :  |     :|: :  |  :| || ||  |: :||  |:  ||
Db         12 AATSFLRAARSGNLDKALDHLRNGVDINTCNQKG--NTALHIAALAGQDEVVRELVNYGA 69

Qy        157 NVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIV 216
              || |:   |  ||: |    | |||  || :||: |      :|||  |  :|  :|
Db         70 NVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAH 129

Qy        217 LLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKM 267
              |: :|  :  :         || | |||          |||      ::|  :::|
Db        130 LINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG----- 173

Qy        268 MALLTPLNVNCHASDGRKS-----------------TPLHLAAGYNRVKIVQLLLQHGA 309
                 |||::  |  :   :                 ||||:|:    | :|:|||  ||
Db        174 ---FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGA 230

Qy        310 DVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSL 369
               :  | | :| ||| |   ||  ::|:|: ||| : |      :|:| ||  : ::   |
Db        231 QIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRL 290

Qy        370 LLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEM 429
              || | |:          ||                                  |:|
Db        291 LLQYDAE-----------ID--------------------------DIT---------- 302

Qy        430 VNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDV 489
                  | | || ||   :   :: ::|| |||  | :     ||||:| :| |  |
Db        303 ------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRV 353

Qy        490 VEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-- 547
              :|:::|  | ::|:   | | || |:: |||   : ||  |  ||: :::  | | |
Db        354 MELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAAR 413

Qy        548 ---NENVQQLLQE----------------GISLGNSEADRQLLE-------------- 572
                  |  : |||                   :|::   : |||

Db        414 AGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTP 473

Qy        573 ---AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 628
                 ||: | ||||  |   : |  |   :|   |||| || | :| | | ||:  |  :|
Db        474 LHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVRVAELLLERDAHPNA 531

Qy        629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQ 687
               | || ||| |  : : :: :||:  |   : :  |  :|||| || : : |: : |||
Db        532 AGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQVEVARSLLQ 590

Qy        688 HGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNC 746
              :|     ::  | ||| |  ::|  ::  ||                        ||   | |
Db        591 YGGSANAESVQGVTPLHLAAQEGHAEMVALL-------------------LSKQANGNL 630

Qy        747 RDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLI 806
               :  |   ||||| |   :: ||: |::||  |:|  : |  ||| |: ||:: :   |:
Db        631 GNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHVASHYGNIKLVKFLL 688

Qy        807 KYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VS 860
              :: | |||  |  ::|||:|||:| | :  ||| :|| |   : :| ||| :      :|
Db        689 QHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDGTTPLAIAKRLGYIS 748

Qy        861 ADDVSALLT 869
                ||  ::|
Db        749 VTDVLKVVT 757


RESULT 15
Q13768
ID   Q13768      PRELIMINARY;      PRT;  1719 AA.
AC   Q13768;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alt. ankyrin (Variant 2.2).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hematopoietic;
RX   MEDLINE=90158830; PubMed=2137557;
RA   Lux S.E., John K.M., Bennett V.;
RT   "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT   structure with homology to tissue-differentiation and cell-cycle
RT   control proteins.";
RL   Nature 344:36-42(1990).
DR   EMBL; X16609; CAA34611.1; -.
DR   HSSP; Q00420; 1AWC.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1719 AA;  189011 MW;  F63465D16D975CBF CRC64;

  Query Match           13.6%;  Score 878.5;  DB 4;  Length 1719;
  Best Local Similarity  25.8%;  Pred. No. 1.8e-46;
  Matches  292;  Conservative  157;  Mismatches  375;  Indels  307;  Gaps   28;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
              | : | |: |  |:  | :    |  : :| || ||  |: :||  |:  |||| |:
```

```
Db        52 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 110

Qy       165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224

Db       111 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 170

Qy       225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275

Db       171 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 211

Qy       276 VNCHASDGRKS-----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317

Db       212 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETKTKD 271

Qy       318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377

Db       272 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYDAE- 330

Qy       378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437

Db       331 ----------ID-------------------------DIT----------------LD 337

Qy       438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497

Db       338 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRVMELLLKTG 394

Qy       498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQ 552

Db       395 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 454

Qy       553 QLLQE-----------------GISLGNSEADRQLLE-----------------AAKAG 577

Db       455 YLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAREG 514

Qy       578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636

Db       515 HVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVRVAELLLERDAHPNAAGKNGLTP 572

Qy       637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695

Db       573 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQVEVARSLLQYGGSANAE 631

Qy       696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754

Db       632 SVQGVTPLHLAAQEGHAEMVALL-------------------LSKQANGNLGNKSG--L 669

Qy       755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814

Db       670 TPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 729

Qy       815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL--------------VS 860

Db       730 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDGTTPLAIAKRLGYISVTDVLKVV 789

Qy       861 ADDVSALLTAAMPPSALPSC-----------------------------------YKPQ 884

Db       790 TDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISFKAERRDSRDVDEEKELLDFVPK 849

Qy       885 VLNGVRSPGATADALSSGPSSPSSLS-----AASSLDNLSGSFSELSSVVSSSGTEGASS 939

Db       850 LDQVVESPAI--------PRIPCAMPETVVIRSEEQEQASKEYDEDSLIPSSPATETSDN 901

Qy       940 LEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHK 999

Db       902 ISPVASPVHTGFLVSFM----------------------------VDARGGSMRGSRHN 932

Qy      1000 LIKGVE-------------RLISGQQGLNPYLTLNTSGSGTILIDLSP------------ 1034

Db       933 GLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEGLASRIIALGPTGAQFLSPVIVE 992

Qy      1035 ---------DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 1076

Db       993 IPHFASHGRGDRELVVLRSENGSVWKEHR---------SRYGESYLDQILN 1034
```

```
RESULT 16
Q61302
ID   Q61302      PRELIMINARY;      PRT;  1848 AA.
AC   Q61302;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Erythroid ankyrin.
GN   ANK1 OR ANK-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Nervous system;
RX   MEDLINE=93252825; PubMed=8486643;
RA   Birkenmeier C.S., White R.A., Peters L.L., Hall E.J., Lux S.E.,
RA   Barker J.E.;
RT   "Complex patterns of sequence variation and multiple 5' and 3' ends
RT   are found among transcripts of the erythroid ankyrin gene.";
RL   J. Biol. Chem. 268:9533-9540(1993).
DR   EMBL; X69063; CAA48801.1; -.
DR   PIR; S37771; S37771.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:88024; Ank1.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1848 AA;  202577 MW;  5744BECBF9EBA056 CRC64;

  Query Match             13.6%;  Score 876;  DB 11;  Length 1848;
  Best Local Similarity   29.7%;  Pred. No. 3e-46;
  Matches  250;  Conservative  132;  Mismatches  281;  Indels  178;  Gaps   21;

Qy       105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164

Db        56 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 114

Qy       165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224

Db       115 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 174

Qy       225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275

Db       175 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 215

Qy       276 VNCHASDGRKS-----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317

Db       216 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETRTKD 275

Qy       318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377

Db       276 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYNAE- 334

Qy       378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437

Db       335 ----------ID-------------------------DIT----------------LD 341
```

```
Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
Db        342 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHIRVMELLLKTG 398

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM----GNENVQQ 553
Db        399 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 458

Qy        554 LLQEGISLGNSEAD------------------RQLLE------------------AAKAG 577
Db        459 YLLQNKAKANAKADDQTPLHCAARIGHTGMVKLLLENGASPNLATTAGHTPLHTAAREG 518

Qy        578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
Db        519 HVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVRLAELLLEHDAHPNAAGKNGLTP 576

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
Db        577 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQIEVARSLLQYGGSANAE 635

Qy        696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
Db        636 SVQGVTPLHLAAQEGHTEMVALL--------------------LSKQANGNLGNKSG--L 673

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
Db        674 TPLHLVSQEGHVPVADVLIKHGVTVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 733

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALL 868
Db        734 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNGTTPLAIAKRLGYISVTDVLKVV 793

Qy        869 T 869
Db        794 T 794
```

```
RESULT 17
Q8C8R3
ID   Q8C8R3      PRELIMINARY;      PRT;  1219 AA.
AC   Q8C8R3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ankyrin 2 (Fragment).
GN   ANK2 OR AI835472.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Retina;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK044634; BAC32012.1; -.
DR   MGD; MGI:88025; Ank2.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 23.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
FT   NON_TER    1219   1219
SQ   SEQUENCE   1219 AA;  131201 MW;  AFBD462967AD6184 CRC64;
```

```
  Query Match           13.5%;  Score 872;  DB 11;  Length 1219;
  Best Local Similarity  29.5%;  Pred. No. 2.8e-46;
  Matches  281;  Conservative  147;  Mismatches  383;  Indels  142;  Gaps   24;

Qy         84 GAACASAAAEAV--EPAARELFEACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAA 140
Db         11 GGACTPQNPPRIRQSDSNASFLRAARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAA 68

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
Db         69 KEGHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFT 128

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL- 255
Db        129 PLYMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVR 188

Qy        256 ---LESARSGNEEKMMALLTPLNVNCHASDGRKS-----------TPLHLAAGYNRVKIV 301
Db        189 LPALHIAARKDDTKSAALLLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVA 245

Qy        302 QLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASK 361
Db        246 TLLLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARS 305

Qy        362 NRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRI 421
Db        306 GHDQVVELLLERKA-PLLARTKN--------------------GLSPLHMAAQGDHVEC 343

Qy        422 KKHLSLEMVNFKHP-----QTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 476
Db        344 VKHL----LQYKAPVDDVTLDYLTALHVAA---HCGHYRVTKLLLDKRANPNARALNGFT 396

Qy        477 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNII 536
Db        397 PLHIACKKNRIKVMELLVKYGASIQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVT 456

Qy        537 SLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDI 596
Db        457 NIRGETALHM-------------------------AARAGQVEVVR--CLLRNGALVDA 488

Qy        597 EGR-QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 655
Db        489 RAREEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTNGYTPLHISAREGQVDVASVLLEA 548

Qy        656 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD 715
Db        549 GAAHSLATKKGFTPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHVAAHYDN---- 604

Qy        716 LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQH 775
Db        605 -----------QKVALLLLEKGASP-----HATAKNGYTPLHIAAKKNQMQIASTLLNY 647

Qy        776 GADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLC 835
Db        648 GAETNTVTKQGVTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGLTSLHLAAQEDKVNVA 707

Qy        836 ALLLAHGADPTLKNQEGQTPLDLV----SADDVSALLTAAMPPSA-LPSCYKP------- 883
Db        708 DILTKHGADRDAYTKLGYTPLIVACHYGNVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQ 767

Qy        884 ---QVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSL 940
Db        768 GHTHIINVLLQHGAKPNATTA--NGNTALAIAKRLGYI--SVVDTLKVVTEEVTTTTTTI 823

Qy        941 EKKEVPGVDFSITQFVRNLGLEHLMDIFERE---QITLDVLVEMGHKELKEIG 990
Db        824 TEKHKLNVPETMTE---------VLDVSDEEGDDTVTGDGGEYLRPEDLKELG 867
```

```
RESULT 18
Q9V4B1
ID   Q9V4B1      PRELIMINARY;      PRT;  1549 AA.
AC   Q9V4B1;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ank protein.
GN   ANK OR CG1651 OR CG32008.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003845; AAF59369.2; -.
DR   HSSP; P80144; 2MYO.
DR   FlyBase; FBgn0011747; Ank.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0045170; C:spectrosome; IDA.
DR   GO; GO:0007016; P:cytoskeletal anchoring; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1549 AA;  170167 MW;  3ECEC852B1484DFF CRC64;
```

```
Query Match           13.4%;  Score 868.5;  DB 5;  Length 1549;
Best Local Similarity  27.4%;  Pred. No. 6.7e-46;
Matches  311;  Conservative  174;  Mismatches  491;  Indels  159;  Gaps   34;

Qy        98 AARELFEACRNGDVERVKRLV---TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQN 154
Db        38 ATISFLRAARSGDIKKVMDFLDCGEISDINSCNANGLNA--LHLAAKDGYVDICCELLRR 95

Qy       155 GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVC 214
Db        96 GIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

Qy       215 IVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDELLESARSGNEEK 266
Db       156 RTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLENDVRGKVRLPALHIAAKKNDVN 215

Qy       267 MMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNAC 326
Db       216 AAKLLLQHDPNADIVSKSGFTPLHIAAHYGNVDIATLLLNNKADVNYVAKHNITPLHVAC 275

Qy       327 SYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN-K 385
Db       276 KWGKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLLQQNA-PILTKTKNGL 334

Qy       386 SAIDLAPTPQLKERLAYEPKGHSLLQAAREADVTRIKKHLSLE----------MVNF 432
Db       335 SALHMAAQGE------HDEAAHLLLDNKAPVDEVTVDYLTALHVAAHCGHVKVAKLLLDY 388

Qy       433 K-HPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 487
Db       389 KANPNARALNGFTPLHIACKK---NRIKMVELLIKHGANIGATTESGLTPLHVASFMGCI 445

Qy       488 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL--- 544
Db       446 NIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDAIAREGQTPLHVA 504

Qy       545 -QMGNEN-VQQLLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db       505 SRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENIVQVLLENGAEN-NAVTKKG 563

Qy       601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db       564 FTPLHLACKYGKQNVVQILLQNGASIDFQGKNDVTPLHVATHYNNPSIVELLLKNGSSPN 623

Qy       661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRG 719
Db       624 LCARNGQCAIHIACKKNYLEIAMQLLQHGADVNIISKSGFSPLHLAAQGGNVDMVQLLLE 683

Qy       720 DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 779
Db       684 YGVISAAAKNGL----------------------TPLHVAAQEGHVLVSQILLEHGANI 720

Qy       780 NAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLL 839
Db       721 SERTRNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAAQQGHIMIINLLL 780

Qy       840 AHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG----------V 889
Db       781 RHKANPNALTKDGNTALHIASN------LGYVTVMESLKIVTSTSVINSNIGAIEEKLKV 834

Qy       890 RSPGATADALSSGPSSPSSLSAAS-------SLDNLSGSFSELSSVVSSSGTE----GAS 938
Db       835 MTPELMQETLLSDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFDTTNTDHDLTDVS 894

Qy       939 SLEKKEVPGVDFSITQFVRNLGLEHLMD--IFEREQITLDVLVEMGHKELKEIGINAYGH 996
Db       895 VLNKKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF---LVDARGGSMRGY 948

Qy       997 RHKLIKGVE------------RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS-- 1041
Db       949 RHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSRILEMSPVDGMFLSPI 1008
```

```
Qy       1042 -VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
Db       1009 TLEVPHYGTLRK----------NEREII-ILRSDNGESWREH-NLYKDIIGEDINQTEE 1055

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIG--GGT 1153
Db       1056 ---FHSDRIVR-IVTQNVPHFFAVV----------SRVRQEVHVIGPDGGT 1092


RESULT 19
Q24241
ID   Q24241      PRELIMINARY;      PRT;  1549 AA.
AC   Q24241;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ankyrin.
GN   ANK OR CG1651 OR CG32008.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A631;
RX   MEDLINE=95024098; PubMed=7937942;
RA   Dubreuil R.R., Yu J.;
RT   "Ankyrin and beta-spectrin accumulate independently of alpha-spectrin
RT   in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:10285-10289(1994).
DR   EMBL; L35601; AAC37208.1; -.
DR   PIR; T13940; T13940.
DR   HSSP; P80144; 2MYO.
DR   FlyBase; FBgn0011747; Ank.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0045170; C:spectrosome; IDA.
DR   GO; GO:0007016; P:cytoskeletal anchoring; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1549 AA;  170096 MW;  587FB6BF88E657CE CRC64;

  Query Match             13.4%;  Score 866.5;  DB 5;  Length 1549;
  Best Local Similarity   27.4%;  Pred. No. 9e-46;
  Matches  311;  Conservative  173;  Mismatches  492;  Indels  159;  Gaps   34;

Qy         98 AARELFEACRNGDVERVKRLV---TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQN 154
Db         38 ATISFLRAARSGDIKKVMDFLDCGEISDINSCNANGLNA--LHLAAKDGYVDICCELLRR 95

Qy        155 GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVC 214
Db         96 GIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

Qy        215 IVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDELLESARSGNEEK 266
Db        156 RTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLENDVRGKVRLPALHIAAKKNDVN 215

Qy        267 MMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNAC 326
Db        216 AAKLLLQHDPNADIVSKSGFTPLHIAAHYGNVDIATLLLNNKADVNYVAKHNITPLHVAC 275

Qy        327 SYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN-K 385
Db        276 KWGKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLLQQNA-PILTKTKNGL 334

Qy        386 SAIDLAPTPQLKERLAYEPKGHSLLQAAREADVTRIKKHLSLE------------MVNF 432
Db        335 SALHMAAQGE------HDEAAHLLLDNKAPVDEVTVDYLTALHVAAHCGHVKVAKLLLDY 388

Qy        433 K-HPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 487
Db        389 KANPNARALNGFTPLHIACKK---NRIKMVELLIKHGANIGATTESGLTPLHVASFMGCI 445

Qy        488 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL--- 544
Db        446 NIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDAIVREGQTPLHVA 504

Qy        545 -QMGNEN-VQQLLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        505 SRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENIVQVLLENGAEN-NAVTKKG 563

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        564 FTPLHLACKYGKQNVVQILLQNGASIDFQGKNDVTPLHVATHYNNPSIVELLLKNGSSPN 623

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRG 719
Db        624 LCARNGQCAIHIACKKNYLEIAMQLLQHGADVNIISKSGFSPLHLAAQGGNVDMVQLLLE 683

Qy        720 DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 779
Db        684 YGVISAAAKNGL----------TPLHVAAQEGHVLVSQILLEHGANI 720

Qy        780 NAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLL 839
Db        721 SERTRNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAAQQGHIMIINLLL 780

Qy        840 AHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG----------V 889
Db        781 RHKANPNALTKDGNTALHIASN------LGYVTVMESLKIVTSTSVINSNIGAIEEKLKV 834

Qy        890 RSPGATADALSSGPSSPSSLSAAS-------SLDNLSGSFSELSSVVSSSGTE----GAS 938
Db        835 MTPELMQETLLSDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFDTTNTDHDLTDVS 894

Qy        939 SLEKKEVPGVDFSITQFVRNLGLEHLMD--IFEREQITLDVLVEMGHKELKEIGINAYGH 996
Db        895 VLNKKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF---LVDARGGSMRGY 948

Qy        997 RHKLIKGVE----------RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS-- 1041
Db        949 RHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSRILEMSPVDGMFLSPI 1008

Qy       1042 -VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
Db       1009 TLEVPHYGTLRK----------NEREII-ILRSDNGESWREH-NLYKDIIGEDINQTEE 1055

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIG--GGT 1153
Db       1056 ---FHSDRIVR-IVTQNVPHFFAVV----------SRVRQEVHVIGPDGGT 1092


RESULT 20
P97582
ID   P97582      PRELIMINARY;      PRT;   843 AA.
AC   P97582;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ankyrin (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

```
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Brain;
RA   Cosentino M.T., Jones O.T.;
RT   "Rat brain ankyrin membrane binding domain.";
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U65916; AAB47551.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 23.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER       1      1
FT   NON_TER     843    843
SQ   SEQUENCE   843 AA;  89982 MW;  DC5A6AC78825D223 CRC64;

  Query Match             13.3%;  Score 858;  DB 11;  Length 843;
  Best Local Similarity   29.6%;  Pred. No. 1.2e-45;
  Matches  275;  Conservative  148;  Mismatches  367;  Indels  140;  Gaps   23;

Qy        105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
              | | |:::: |: |      :|: :  |  :  || ||  |   :|: ||  |::| :
Db          3 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 60

Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
              |   || |   | |||| :|:: ||: ||:    :|||: || :  |||   ||::||
Db         61 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 120

Qy        224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
              :    || |    ||       | |:|     | ::    | |   :: |  |||
Db        121 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 178

Qy        276 VNCHASDGRKS-----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
              | | :| :              ||||:|| |  | :  |||  || |    :  : |||
Db        179 -NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 237

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
              |    |:  : :||:  |  ::|      ||||  ||      :|  |||  || | |     |
Db        238 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 296

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
                                  | | |  | : |      |||    :  | |       :
Db        297 ---------------------GLSPLHMAAQGDHVECVKHL----LQHKAPVDDVTLDYL 331

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              |||| ||    :    :: :||| | || | :      ||||:| :|     |:|::||: |
Db        332 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAY 388

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |:   | | :  ||: |||     |||  |  |:: ::: | ||| |
Db        389 IQAITESGLTPIPVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHM------------ 435

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
                        ||:||:|| |:  | :::    |  | : |||| |:   :  :|:
Db        436 ------------AARAGEVEVVR--CLLRNGALVDARAREEQTPLHIASRLGKTEIVQL 480

Qy        619 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 678
              |||| |    |    | ||| :   |  :|| :|:: ||  ::|    ||||| ||  |
Db        481 LLQHMAHPDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATKKGFTPLHVAAKYGS 540

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              :: |||||  |       ::| ||| :    |                    :|  |  ::|
Db        541 LDVAKLLLQRRAAADSAGKNGLTPLHVAAHYDN----------------QKVALLLLEKG 584

Qy        739 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 798
              :||       |    ||||:||  | :::|  || :||: |   | |: ||| |:  ||
Db        585 ASP-----HATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNTVTKQGVTPLHLASQEGH 639

Qy        799 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 858
```

```
                |:  |:::  | :: : |   | || ||:: :  :  :|  ||||     : | ||| :
Db        640 TDMVTLVLEKGANIHMSTKSGLTSLHLAAEEDKVNVADILTKHGADQDAYTKLGYTPLIV 699

Qy        859 V----SADDVSALLTAAMPPSA-LPSCYKP----------QVLNGVRSPGATADALSSGP 903
                  :   |: ||      :|   : | |           ::| :   ||  :| ::
Db        700 ACHYGNVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTA-- 757

Qy        904 SSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEH 963
              :  ::|: |  |  :  |  :   ||:   |   ::: :|       ::|:
Db        758 NGNTALAIAKRLGYI--SVVDTLKVVTEEVTTTTTTITEKHKLNAPETMTE--------- 806

Qy        964 LMDIFERE---QITLDVLVEMGHKELKEIG 990
              ::|: : |    :| |     :  ::|||:|
Db        807 VLDVSDEEGDDTVTGDGGEYLRPEDLKELG 836


RESULT 21
Q17487
ID   Q17487      PRELIMINARY;      PRT;  1809 AA.
AC   Q17487;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   E. elegans ankyrin-related unc-44 (GB:U21734).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gattung S.;
RT   "The sequence of C. elegans cosmid B0350.";
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAA93446.1; -.
DR   HSSP; P42773; 1IHB.
DR   WormPep; B0350.2b; CE06703.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
```

```
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1809 AA;  198193 MW;  AD86349AA64534F8 CRC64;

  Query Match           13.3%;  Score 858;  DB 5;  Length 1809;
  Best Local Similarity   26.3%;  Pred. No. 4e-45;
  Matches  319;  Conservative  163;  Mismatches  443;  Indels  286;  Gaps   39;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA-------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL----------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV----------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG--- 1024
Db        839 FSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGGAIP 898

Qy       1025 ---SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQKVC 1075
Db        899 KINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLHGAG 953

Qy       1076 NKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DERHA 1123
Db        954 PEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDARGG 1013

Qy       1124 YIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQ 1183
Db        1014 AMRGCRHSGVRIIVPPRKASQ----------PIR-----------VTCRY-LRKD--- 1046

Qy       1184 FSAMKMAHSPP 1194
Db        1047 ----KLAHPPP 1053
```

RESULT 22
Q8MQG0
ID   Q8MQG0   PRELIMINARY;   PRT; 1841 AA.
AC   Q8MQG0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   UNC-44 protein (corresponding sequence B0350.2e).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Gattung S.;
RT   "The sequence of C. elegans cosmid B0350.";
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAM75382.1; -.
DR   WormPep; B0350.2e; CE31147.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 23.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.

```
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1841 AA;  201234 MW;  24CC6020DDCE0FC8 CRC64;

  Query Match           13.3%;  Score 858;  DB 5;  Length 1841;
  Best Local Similarity   26.3%;  Pred. No. 4.1e-45;
  Matches  319;  Conservative  163;  Mismatches  443;  Indels  286;  Gaps   39;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL----------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV----------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG--- 1024
```

```
Db        839 FSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGGAIP 898

Qy       1025 ---SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQKVC 1075
Db        899 KINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLHGAG 953

Qy       1076 NKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DERHA 1123
Db        954 PEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDARGG 1013

Qy       1124 YIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQ 1183
Db       1014 AMRGCRHSGVRIIVPPRKASQ----------PIR-----------VTCRY-LRKD--- 1046

Qy       1184 FSAMKMAHSPP 1194
Db       1047 ----KLAHPPP 1053
```

```
RESULT 23
Q17486
ID   Q17486      PRELIMINARY;      PRT;  1867 AA.
AC   Q17486; O02516;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   C. ELEGANS ankyrin-related UNC-44 (GB:U21734).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gattung S.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAA93444.1; -.
DR   EMBL; U39847; AAB41828.1; -.
DR   HSSP; P42773; 1IHB.
DR   WormPep; B0350.2c; CE06704.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1867 AA;  204253 MW;  4689A2104623B4C6 CRC64;

  Query Match           13.3%;  Score 858;  DB 5;  Length 1867;
  Best Local Similarity  26.3%;  Pred. No. 4.2e-45;
  Matches  319;  Conservative  163;  Mismatches  443;  Indels  286;  Gaps   39;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL----------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV----------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
```

```
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD-------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG--- 1024
Db        839 FSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGGAIP 898

Qy       1025 ---SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQKVC 1075
Db        899 KINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLHGAG 953

Qy       1076 NKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DERHA 1123
Db        954 PEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDARGG 1013

Qy       1124 YIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQ 1183
Db       1014 AMRGCRHSGVRIIVPPRKASQ----------PIR-----------VTCRY-LRKD--- 1046

Qy       1184 FSAMKMAHSPP 1194
Db       1047 ----KLAHPPP 1053
```

```
RESULT 24
Q17489
ID   Q17489      PRELIMINARY;      PRT;  2039 AA.
AC   Q17489;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   C. elegans ankyrin-related unc-44 (GB:U21734).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gattung S.;
RT   "The sequence of C. elegans cosmid B0350.";
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAA93443.1; -.
DR   PIR; T15347; T15347.
DR   HSSP; P42773; 1IHB.
DR   WormPep; B0350.2a; CE06702.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   2039 AA;  222850 MW;  7722C4C2E6AFD68A CRC64;

  Query Match           13.3%;  Score 858;  DB 5;  Length 2039;
  Best Local Similarity  26.3%;  Pred. No. 4.8e-45;
  Matches  319;  Conservative  163;  Mismatches  443;  Indels  286;  Gaps   39;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL----------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV---------------RG 458
```

```
Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV-----------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD-------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG--- 1024
Db        839 FSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGGAIP 898

Qy       1025 ---SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQKVC 1075
Db        899 KINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLHGAG 953

Qy       1076 NKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DERHA 1123
Db        954 PEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDARGG 1013

Qy       1124 YIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQ 1183
Db       1014 AMRGCRHSGVRIIVPPRKASQ----------PIR-----------VTCRY-LRKD--- 1046

Qy       1184 FSAMKMAHSPP 1194
Db       1047 ----KLAHPPP 1053
```

```
RESULT 25
Q17488
ID   Q17488      PRELIMINARY;      PRT;  1815 AA.
AC   Q17488; O02517;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   C. ELEGANS ankyrin-related UNC-44 (GB:U21734).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
```

```
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gattung S.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAA93445.1; -.
DR   EMBL; U39847; AAB41826.1; -.
DR   HSSP; P42773; 1IHB.
DR   WormPep; B0350.2d; CE06705.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1815 AA;  198956 MW;  3ECC6E310915C915 CRC64;

  Query Match             13.2%;  Score 854;  DB 5;  Length 1815;
  Best Local Similarity   27.6%;  Pred. No. 7.2e-45;
  Matches  291;  Conservative  146;  Mismatches  387;  Indels  232;  Gaps   30;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265

Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL------------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV----------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLIKGVERLISGQ--QGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE 1043
Db        839 FSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQ-------DSTGVHMIHTG---------- 881

Qy       1044 EEMQSTVREHRDGG-----HAGGIFNRYNILKIQKV 1074
Db        882 EQLLQRSQELENGGAIPKINSGGMSPEKEFAKIAPV 917


RESULT 26
Q17343
ID   Q17343      PRELIMINARY;      PRT;  6994 AA.
AC   Q17343;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   UNC-44 ankyrins.
GN   UNC-44.
OS   Caenorhabditis elegans.
```

```
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
RN   [2]
RP   REVISIONS, AND SEQUENCE OF 6126-6994 FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U39847; AAB41827.1; -.
DR   EMBL; U21733; AAB38384.1; -.
DR   HSSP; P42773; 1IHB.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   6994 AA;  775364 MW;  90CB449925D9923D CRC64;

  Query Match             13.2%;  Score 854;  DB 5;  Length 6994;
  Best Local Similarity   27.6%;  Pred. No. 6e-44;
  Matches  291;  Conservative  146;  Mismatches  387;  Indels  232;  Gaps   30;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA-------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL-----------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV----------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD-------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLIKGVERLISGQ--QGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE 1043
Db        839 FSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQ-------DSTGVHMIHTG---------- 881

Qy       1044 EEMQSTVREHRDGG-----HAGGIFNRYNILKIQKV 1074
Db        882 EQLLQRSQELENGGAIPKINSGGMSPEKEFAKIAPV 917


RESULT 27
Q17490
ID   Q17490      PRELIMINARY;      PRT;  6994 AA.
AC   Q17490;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Uncoordinated protein 44, isoform f.
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Gattung S.;
RT   "The sequence of C. elegans cosmid B0350.";
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
```

```
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAA93447.2; -.
DR   PIR; T15348; T15348.
DR   WormPep; B0350.2f; CE31847.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 23.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
SQ   SEQUENCE   6994 AA;  775382 MW;  0FFBBEEAD9408975 CRC64;

  Query Match           13.2%;  Score 854;  DB 5;  Length 6994;
  Best Local Similarity   27.6%;  Pred. No. 6e-44;
  Matches  291;  Conservative  146;  Mismatches  387;  Indels  232;  Gaps   30;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA-------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL----------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV---------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
```

```
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 992
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        993 -------AYGHRHKLIKGVERLISGQ--QGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE 1043
Db        839 FSESSEDEGQAAEHEAVAAHEKDFSDNLTQGLQ-------DSTGVHMIHTG---------- 881

Qy       1044 EEMQSTVREHRDGG-----HAGGIFNRYNILKIQKV 1074
Db        882 EQLLQRSQELENGGAIPKINSGGMSPEKEFAKIAPV 917
```

```
RESULT 28
Q17344
ID   Q17344      PRELIMINARY;      PRT;  1786 AA.
AC   Q17344;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   UNC-44 (Fragment).
GN   UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
DR   EMBL; U21734; AAA85854.1; -.
DR   HSSP; P42773; 1IHB.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
```

```
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
FT   NON_TER    1786   1786
SQ   SEQUENCE   1786 AA;  195550 MW;  03E220FB521747F1 CRC64;

  Query Match             13.2%;  Score 851;  DB 5;  Length 1786;
  Best Local Similarity   26.3%;  Pred. No. 1.1e-44;
  Matches  320;  Conservative  161;  Mismatches  444;  Indels  292;  Gaps   38;

Qy         91 AAEAVEPAARE----LPEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205
Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265
Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444
Db        291 KDLL-------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502
Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550
Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV----------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776
Db        579 RGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673

Qy        837 LLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPSCYK 882
Db        674 LLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-ACHF 732

Qy        883 PQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        733 GQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NGASP 781

Qy        940 LEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGINAY- 994
Db        782 NEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMNETM 838

Qy        995 ---------------GHRHKL------------IKGVERLISGQQGLNPYLTLNTSG--- 1024
Db        839 FSESEDEGQAAEHEGGGAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGGAIP 898

Qy       1025 ---SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQKVC 1075
Db        899 KINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLHGAG 953

Qy       1076 NKKLWERYTHRRKE------------VSEENHNHANERMLFHGSPFVNAIIHKGF----- 1118
Db        954 PEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSKLQHRTFLISFL 1013

Qy       1119 -DERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTL 1177
Db        1014 VDARGGAMRGCRHSGVRIIVPPRKASQ----------PIR------------VTCRY-L 1049

Qy       1178 GKSFLQFSAMKMAHSPP 1194
Db       1050 RKD-------KLAHPPP 1059


RESULT 29
Q9NCP8
ID   Q9NCP8      PRELIMINARY;      PRT;  1159 AA.
AC   Q9NCP8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ankyrin 2 (CG7462-PC).
GN   ANK2 OR CG7462.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20304926; PubMed=10844021;
RA   Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA   Hortsch M.;
RT   "The L1-type cell adhesion molecule neuroglian influences the
RT   stability of neural ankyrin in the Drosophila embryo but not its
RT   axonal localization.";
RL   J. Neurosci. 20:4515-4523(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
```

```
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF190635; AAF73309.1; -.
DR   EMBL; AE003558; AAN12046.1; -.
DR   HSSP; P42773; 1IHB.
DR   FlyBase; FBgn0017645; Ank2.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1159 AA;  125769 MW;  9ED146C9E483DE58 CRC64;

  Query Match           13.1%;  Score 848.5;  DB 5;  Length 1159;
  Best Local Similarity  28.1%;  Pred. No. 7.9e-45;
  Matches  280;  Conservative  146;  Mismatches  389;  Indels  183;  Gaps   23;

Qy       105 ACRNGDVERV-KRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
Db        18 AARAGNLERVLEHLKNNIDINTSNANGLNA--LHLASKDGHIHVVSELLRRGAIVDSATK 75

Qy       164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223

Db        76 KGNTALHIASLAGQEEVVKLLLEHNASVNVQSQNGFTPLYMAAQENHDAVVRLLLSNGAN 135

Qy       224 PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDG 283
Db       136 QSLATEDGFT------PLAVAMQQG--------------HDKVVAVLLESDTR-----G 169

Qy       284 RKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGA 342
Db       170 KVRLPALHIAAKKDDVKAATLLLDNDHNPDVTSKSGFTPLHIASHYGNQNIANLLIQKGA 229

Qy       343 CVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAY 402
Db       230 DVNYSAKHNISPLHVAAKWGKTNMVSLLLEKGG-----NIEAKTRDGLTP---------- 274

Qy       403 EFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLR 462
Db       275 ------------------------------------LHCAARSGH---EQVVDMLLE 292

Qy       463 KGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQT 522
Db       293 RGAPISAKTKNGLAPLHMAAQGEHVDAARILLYHRAPVDEVTVDYLTALHVAAHCGHVRV 352

Qy       523 CRLLLSYGCDPNIISLQGFTALQMGNEN-----VQQLLQEGIS---------------- 560
Db       353 AKLLLDRNADANARALNGFTPLHIACKKNRLKVVELLLRHGASISATTESGLTPLHVAAF 412

Qy       561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
Db       413 MGCMNIVIYLLQHDASPDVPTV---------------RGETPLHLAARANQTDIIRILL 456

Qy       621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
Db       457 RNGAQVDARAREQQTPLHIASRLGNVDIVMLLLQHGAQVDATTKDMYTALHIAAKEGQDE 516

Qy       681 ICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKG--------- 730
Db       517 VAAVLIENGAALDAATKKGFTPLHLTAKYGHIKVAQLLLQKEADVDAQGKNGVTPLHVAC 576

Qy       731 -------CLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQD 783
Db       577 HYNNQQVALLLLEKGASP-----HATAKNGHTPLHIAARKNQMDIATTLLEYGALANAES 631

Qy       784 KGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGA 843
Db       632 KAGFTPLHLSSQEGHAEISNLLIEHKAAVNHPAKNGLTPMHLCAQEDNVNVAEILEKNGA 691

Qy       844 DPTLKNQEGQTPLDLVS----ADDVSALL-TAAMPPSALPSCYKP----------QVLNG 888
Db       692 NIDMATKAGYTPLHVASHFGQANMVRFLLQNGANVDAATSIGYTPLHQTAQQGHCHIVNL 751

Qy       889 VRSPGATADALSSGPSSP----SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKE 944
Db       752 LLEHKANANAQTVNGQTPLHIARKLGYISVLDSLKTITKEDETAAAPSQAE-----EKYR 806

Qy       945 VPGVDFSITQFVRNL----GLEHLMDIFEREQITLDVLVEMGHKEL-------KEIGINA 993
Db       807 VVAPEAMHESFMSDSEEEGGEDNMLSDQPYRYLTVDEMKSLGDDSLPIDVTRDERMDSNR 866

Qy       994 YGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 1031
Db       867 MTQSAEYASGVPPTI-GEEVISPHKTQVYGSSPKATVD 903


RESULT 30
Q8SWY2
ID   Q8SWY2      PRELIMINARY;      PRT;  1009 AA..
AC   Q8SWY2; Q9VSA2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   GH01626p (CG7462 protein).
GN   ANK2 OR CG7462.
```

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY094999; AAM11327.1; -.
DR EMBL; AE003558; AAF50525.2; -.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 19.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 19.
DR PROSITE; PS50088; ANK_REPEAT; 17.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1009 AA; 109919 MW; 3AD24CEF12A77303 CRC64;

Query Match 12.6%; Score 817; DB 5; Length 1009;
Best Local Similarity 25.9%; Pred. No. 6.4e-43;
Matches 288; Conservative 146; Mismatches 402; Indels 274; Gaps 29;

```
Qy      127 DTAGRKSTP-LHFAAGFGRKDVVE---YLLQNGANVQARDDGGLIPLHNACSFGHAEVVN 182
            || |:   | || ||   :|| |:    || |  |       |  ||| |  :|:  : |
Db       16 DTRGKVRLPALHIAA---KKDDVKAATLLLDNDHNPDVTSKSGFTPLHIASHYGNQNIAN 72

Qy      183 LLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSA 242
            ||:: ||| |     | :||| ||  || ::  :||: |     :  || | |  |
Db       73 LLIQKGADVNYSAKHNISPLHVAAKWGKTNMVSLLLEKGGNIEAKTRDGLTPLHCA---- 128

Qy      243 KAVLTGEYKKDELLESARSGNEEKMMALL---TPLNVNCHASDGRKSTPLHLAAGYNRVK 299
                          ||||:|: :  ||    |::    |      |||:||     |
Db      129 ---------------ARSGHEQVVDMLLERGAPIS----AKTKNGLAPLHMAAQGEHVD 168

Qy      300 IVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAA 359
            ::|| | | |       |  || |  ||  | :||:  |  ||  |  ||||| |
Db      169 AARILLYHRAPVDEVTVDYLTALHVAAHCGHVRVAKLLLDRNADANARALNGFTPLHIAC 228

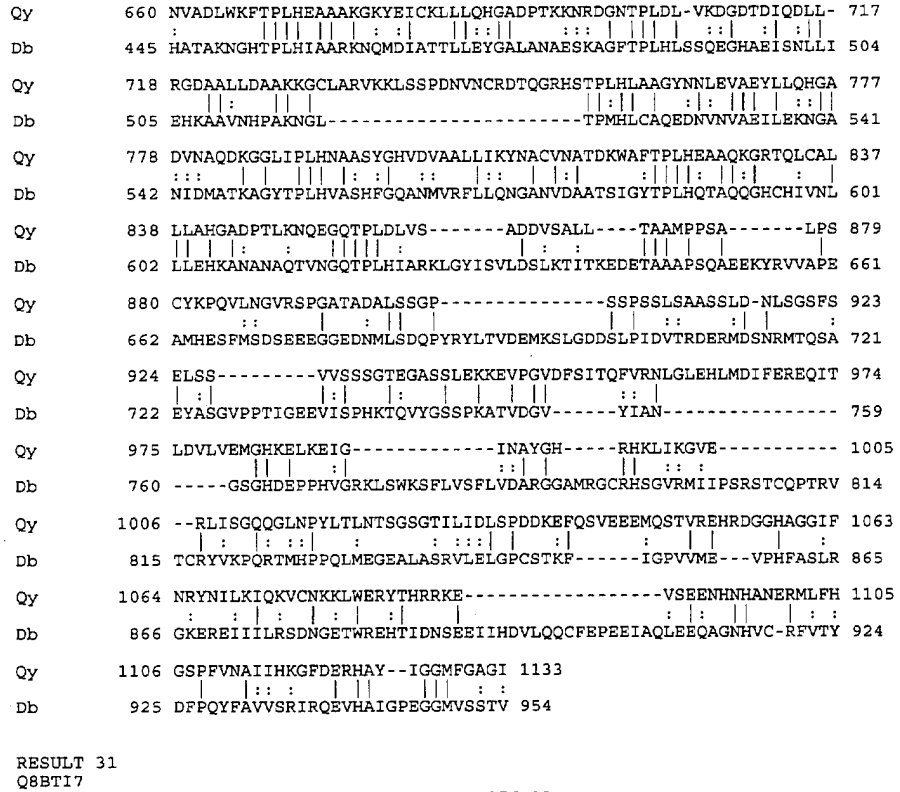
Qy      360 SKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVT 419
            |||::|  ||| :||  :       : : :|       : |      |||      ||
Db      229 KKNRLKVVELLLRHGASISATTESGLTPLHVAAFMGCMNIVIY------LLQHDASPDVP 282

Qy      420 RIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLH 479
            ::                  || || ||  :    :  |  :||| || :: : :|  ||||
Db      283 TVR--------------GETPLHLAARA---NQTDIIRILLRNGAQVDARARERQQTPLH 324

Qy      480 VASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQ 539
            :||   : |:| :::: | |:|:|      |:|| ||  |  :   :|:  |   :  : :
Db      325 IASRLGNVDIVMLLLQHGAQVDATTKDMYTALHIAAKEGQDEVAAVLIENGAALDAATKK 384

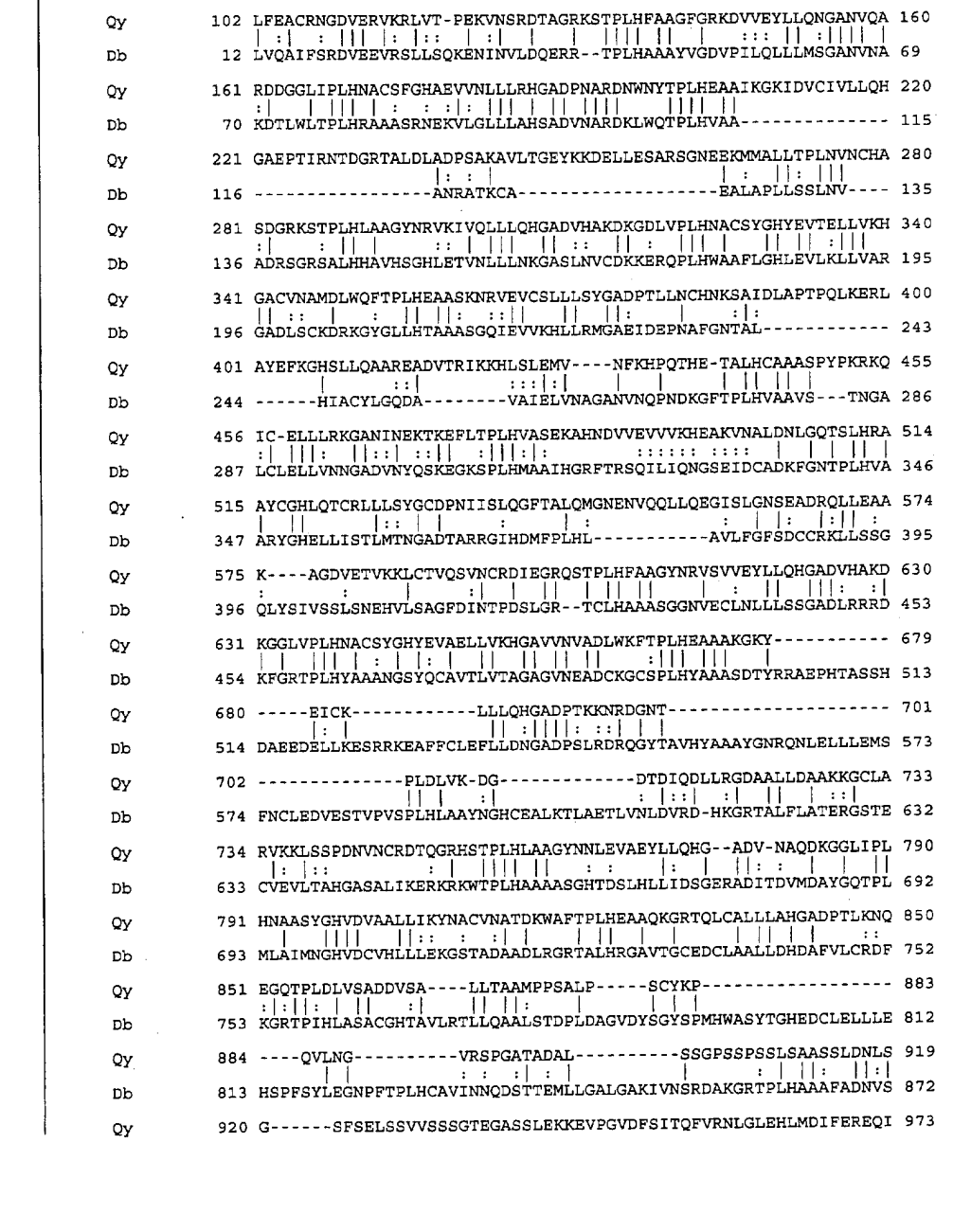
Qy      540 GFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR 599
            ||
Db      385 GF---------------------------------------------------------- 386

Qy      600 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 659
              |||  | |  : | : |||  ||| |: | |: ||| || | : :|| ||:: ||
Db      387 --TPLHLTAKYGHIKVAQLLLQKEADVDAQGKNGVTPLHVACHYNNQQVALLLLEKGASP 444
```



RESULT 31
Q8BTI7
ID Q8BTI7 PRELIMINARY; PRT; 1076 AA.
AC Q8BTI7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical ankyrin-repeat containing protein.
GN G431002C21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK090124; BAC41105.1; -.
DR PIR; PT0649; PT0712.
DR PIR; PT0706; PT0706.
DR MGD; MGI:2444029; G431002C21Rik.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 28.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 28.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein.
SQ SEQUENCE 1076 AA; 115031 MW; 058FE2C95E1A5682 CRC64;

Query Match 11.2%; Score 724.5; DB 11; Length 1076;
Best Local Similarity 26.3%; Pred. No. 5.3e-37;
Matches 287; Conservative 140; Mismatches 386; Indels 279; Gaps 33;



```
Db        873 GLRMLLQHQAEVNATDHTGRTALMTAAESGQTAAVEFLLYR-----GKADLTVLDENKNT 927

Qy        974 TLDVLVEMGHK--------ELKEIG-INA----------YGHRHKLIKGVERLIS----- 1009
Db        928 ALHLACSKGHEKCALMILAETQDLGLINATNSALQMPLHIAARNGLASVVQALLSRGATV 987

Qy       1010 ---GQQGLNPYL 1018
Db        988 LAVDEEGHTPAL 999


RESULT 32
Q8N8A2
ID   Q8N8A2      PRELIMINARY;      PRT;   919 AA.
AC   Q8N8A2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein FLJ39767.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA   Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA   Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA   Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA   Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA   Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA   Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK097086; BAC04946.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 26.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 26.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   919 AA;  99819 MW;  FP98E15F0E9C4E89 CRC64;

  Query Match           11.1%;  Score 714.5;  DB 4;  Length 919;
  Best Local Similarity  28.7%;  Pred. No. 1.8e-36;
  Matches  246;  Conservative  109;  Mismatches  287;  Indels  215;  Gaps   27;

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
Db          7 EDVNTLDS--EKRTPLHVAAFLGDAEIIELLILSGARVNAKDNMWLTPLHRAVASRSEEA 64

Qy        181 VNLLLRHGADPNARD-NWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLAD 239
Db         65 VQVLIKHSADVNARDKNWQ-TPLHVAA---------------------------------- 90

Qy        240 PSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
Db         91 -ANKAV--------------------KCAEVIIPLLSSVNVSDRGGRTALHHAALNGHVE 129

Qy        300 IVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAA 359
Db        130 MVNLLLAKGANINAFDKKDRRALHWAAYMGHLDVVALLINHGAEVTCKDKKGYTPLHAAA 189

Qy        360 SKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVT 419
Db        190 SNGQINVVKHLLNLGVEIDEINVYGNTALHIA---------CYN---------GQDAVVN 231

Qy        420 RIKKHLSLEMVNFKHPQTHE-TALHCAAASPYPKRKQIC-ELLLRKGANINEKTKEFLTP 477
Db        232 ELIDY----GANVNQPNNNGFTPLHFAAASTH---GALCLELLVNNGADVNIQSKDGKSP 284

Qy        478 LHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIS 537
Db        285 LHMTAVHGRFTRSQTLIQNGGEIDCVDKDGNTPLHVAARYGHELLINTLRTSGADTAKCG 344

Qy        538 LQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 597
Db        345 IHSMFPLHLAALN-----------AHSDCCRKLL--SSGFEIDTPDKF------------ 379

Qy        598 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 657
Db        380 GR--TCLHAAAAGGNVECIKLLQSSGADFHKKDKCGRTPLHYAAANCHFHCIETLVTTGA 437

Qy        658 VVNVADLWKFTPLHEAAA--------------------------KGKYEIC-KLLLQHG 689
Db        438 NVNETDDWGRTALHYAAASDMDRNKTILGNAHDNSEELERARELKEKEATLCLEFLLQND 497

Qy        690 ADPTKKNRDG----------------------------------NTPLDLVK-DG---- 709
Db        498 ANPSIRDKEGYNSIHYAAAYGHRQCLELLLERTNSGFEESDSGATKSPLHLAAYNGHHQA 557

Qy        710 ---------DTDIQDLLRGDAALLDAAKKGCLARVKKLSSP-------DNVNCRDTQGRH 753
Db        558 LEVLLQSLVDLDIRD-EKGRTALDLAAFKGHTECVEALINPGASIPVKDNVTKR------ 610

Qy        754 STPLHLAAGYNNLEVAEYLLQHGAD----VNAQDKGGLIPLHNAASYGHVDVAALLIKYN 809
Db        611 -TPLH-ASVINGHTLCLRLLLEIADNPEAVDVKDAKGQTPLMLAVAYGHIDAVSLLLEKE 668

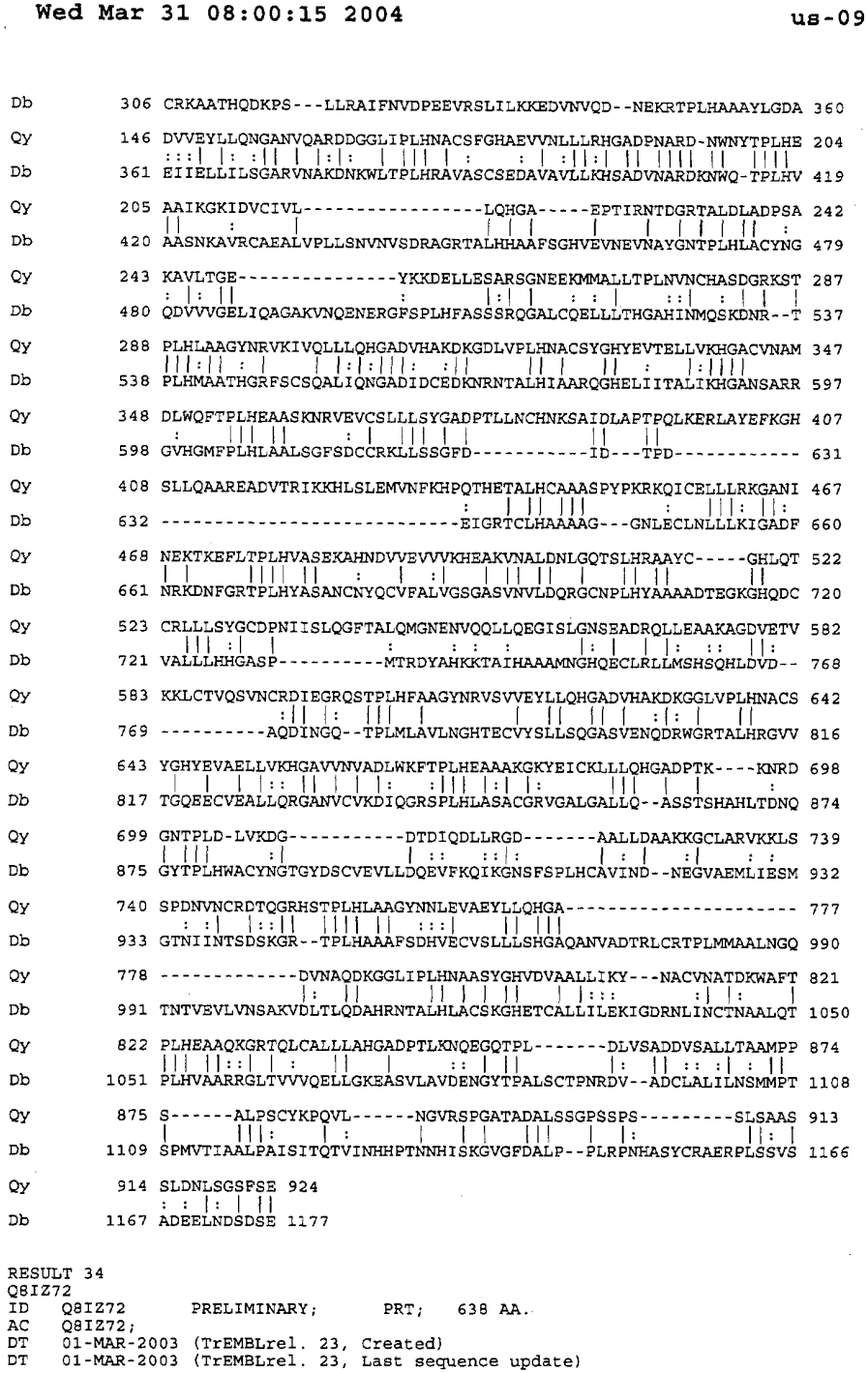
Qy        810 ACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADD----VS 865
Db        669 ANVDTVDILGCTALHRGIMTGHEECVQMLLEQEVSILCKDSRGRTPLHYAAARGHATWLS 728

Qy        866 ALLTAAMPPSALPSCYK 882
Db        729 ELLQMAL--SEEDCCFK 743


RESULT 33
Q8AX92
ID   Q8AX92      PRELIMINARY;      PRT;  1179 AA.
AC   Q8AX92;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kosan C., Kunz J.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF527754; AAO15006.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 23.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   PROSITE; PS50088; ANK_REPEAT; 16.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1179 AA;  127675 MW;  ABFF2F2B2CC14EF2 CRC64;

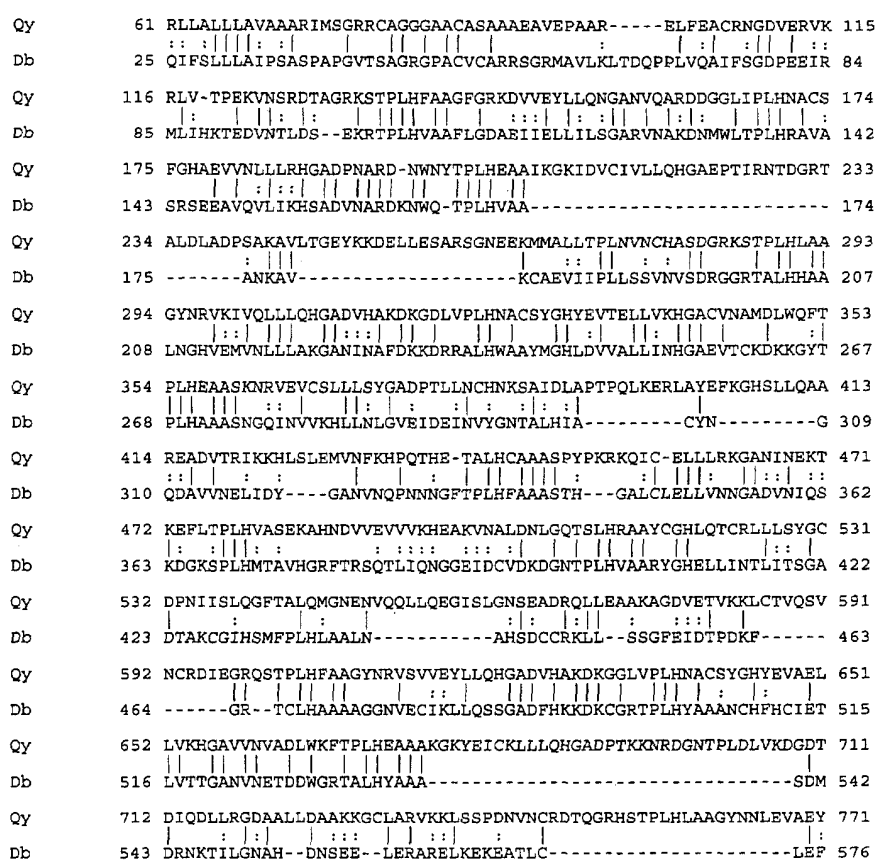
  Query Match           10.9%;  Score 705;  DB 13;  Length 1179;
  Best Local Similarity  28.2%;  Pred. No. 1.1e-35;
  Matches  274;  Conservative  104;  Mismatches  361;  Indels  232;  Gaps   32;

Qy         87 CASAAAEAVEPAARELPEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145
```



RESULT 34
Q8IZ72
ID Q8IZ72 PRELIMINARY; PRT; 638 AA.
AC Q8IZ72;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Duodenum;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC016985; AAH16985.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 16.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 16.
DR PROSITE; PS50088; ANK_REPEAT; 13.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 638 AA; 69275 MW; C1C4E40F7F21B13F CRC64;

Query Match 10.4%; Score 672; DB 4; Length 638;
Best Local Similarity 28.3%; Pred. No. 5e-34;
Matches 219; Conservative 104; Mismatches 272; Indels 178; Gaps 22;



```
Qy        772 LLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLH 824
              |||: |: : :|| |   :| ||:|||      |:  :    : |     || |
Db        577 LLQNDANPSIRDKEGYNSIHYAAAYGHRQCLELVSCF-VLFSLTHTHTHTPNH 628


RESULT 35
Q7T1G6
ID   Q7T1G6      PRELIMINARY;      PRT;  1614 AA.
AC   Q7T1G6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ion channel NompC.
GN   NOMPC.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22728086; PubMed=12805553;
RA   Sidi S., Friedrich R.W., Nicolson T.;
RT   "NompC TRP channel required for vertebrate sensory hair cell
RT   mechanotransduction.";
RL   Science 301:96-99(2003).
DR   EMBL; AY313897; AAP86445.1; -.
SQ   SEQUENCE   1614 AA;  175798 MW;  773343F37E3CB2DD CRC64;

  Query Match             10.2%;  Score 662;  DB 13;  Length 1614;
  Best Local Similarity   24.8%;  Pred. No. 9.3e-33;
  Matches  298;  Conservative  156;  Mismatches  441;  Indels  308;  Gaps   38;

Qy         20 VGAAHG-----------TAPD-PVTAGSQAARALSAS-SPGGLAL---LLAGPGLLLRL-  62
              : |||             | |: |      |      :|| | ||||:   ||       ||
Db         87 IAAAHSKDEIVKLLVRKTDPNSPAGPNDQLPLHYAASRSTGGLAVVQTLLKFSSKDARLT 146

Qy         63 ------LALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKR 116
                  | ||||  |  :  |        :      ||:        |   ||  |||  |
Db        147 PDKNGCLPLLLAAEAGNVGIVRELLSSQSEPQIRAAKTAN-GDTALHICCRRRDVEMAKI 205

Qy        117 LV----TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNA 172
              ||      |:  |          ||||  ||   | ::::::|    ||    |    ||| |
Db        206 LVEFGANPDSQNDEG-----QTPLHIAAHEGDENMLKFLYLCKANANISDKMDRSPLHIA 260

Qy        173 CSFGHAEVVNLLL-RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG 231
                 ||  || :|  :  :   ||      | || |:  |      :  |: |    : |  |
Db        261 AERGHTNVVEILTEKFRSCVLARTKDGNTLLHIASQCGHPTTALSFLRKGVPLHMPNKSG 320

Qy        232 RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTP-LNVVNCHASDGRKSTPLH 290
                |  |                      |: |:   : |||     :|:  | ||:  | ||
Db        321 AVCLHAA-------------------AKRGHTAVVKALLQKGAHVDAAARDGQ--TALH 358

Qy        291 LAAGYNRVKIVQLLLQHGADVHAK-DKGDLVPLHNACSYGHYE-VTELLVKHGACVNAMD 348
              :|    | ::||:|| || |  :  |    ||| :       |   |:|:| || |||
Db        359 IAVENCRPQVVQMLLGFGAHVQLRGGKAQETPLHISARVKEGERAAEMLLKSGAEVNAEQ 418

Qy        349 LWQFTPLHEAASKNRVEVCSLLLSYGADP-------------TLLNCH------------ 383
                | || ||    :::   |: | ||                 : :||
Db        419 ENGETALHVAARHGSLQMIRALIQEGGDPRWRSRVGESPLHVAVRHCHAHVVQEILTFLT 478

Qy        384 -NKSAIDLA-------PTPQLKERLAYEFKGHSLLQ-----------AAREADVTRIKKH 424
               ||  |           :   || | :  :| |                :||:|  : :
Db        479 NEKSRRDAELCVCEGNQDGETALHLAAELRTDALHQPEEDITIIQILMEHQADITAVTR- 537

Qy        425 LSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN-----INEKTKEFLTPLH 479
                      || || || :|         : |:|      |      ||: :|   :||
Db        538 -----------QTGETPLHYSAR--VGNTAVLQEMLRNVPTNQIQTAINKHSKNGWSPLL 584

Qy        480 VASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQ 539
              :|::: | :||:::::: |:|:  |  |: ::| ||  ||      :|||      |  : |
Db        585 LAADQGHTEVVKILLQNNARVDVFDEEGKAAIHLAAQRGHQDIVDVLLSQKAFVNAKTKQ 644
```

```
Qy        540 GFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR 599
              | | | :  :|             |::   | |:|  :|              ||:   : |
Db        645 GLTPLHLSAQN-----------GSARLVRLLVENHQA------------SVDALSL--R 678

Qy        600 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVK----- 654
              : |||| ||   :: |   ||   ||: | |  |  |||  |     | || :| ::
Db        679 KQTPLHLAAMSGQLDVCSSLLNLRADITATDSRGQTPLHLAAESDHSEVVKLFLRLRPEL 738

Qy        655 -----------------HGAVVNVADLWKFT---------------PLHEAAAKGKYEIC 682
                              |:|  : :|  |                  |||  |||  |  |:
Db        739 STLANEDGSTCTHIAAAKGSVSVIRELLMFNQGGVGTLNHKAHGLCPLHLAAAGGHAEVV 798

Qy        683 KLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAAL-LDAAKKGCLA------- 733
              |:||: ||  |::: :| | : |  | | | | ::|||    | : ::| |  |
Db        799 KVLLEAGASVTEEDAEGMTAVHLAAKHGHTHILEVLRGSVPLKIQSSKTGFTALHVAASF 858

Qy        734 -------------------RVKKLSSPDNV------------------------------ 744
                                :|  |::
Db        859 GQMNFVREILTKVPATIRSEFPTISGKDDIKRQQPLAESGFTPLHLASQSGHESVVRLLL 918

Qy        745 NC------RDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ-DKGGLIPLHNAASYG 797
              ||       :|  : |:||||||   :  |  ||   : :  | |: |    || ||::|
Db        919 NCPGVQADAETNIQGSSPLHLAAQSGHTAVVGLLLSRSSSLLHQADRRGRSALHLAAAHG 978

Qy        798 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 857
              |||:  :|:   | :| ||   :| || ||: |  ::   |:  ||    :  |:|||
Db        979 HVDMVRVLLGQGAEINHTDMSGWTALHYAAEAGCLEVLLFLVESGASACAECHGGRTPLQ 1038

Qy        858 LVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDN 917
                :  :                ::  |:  :|     |    |       : :      ||
Db       1039 YAAQQN----------------HESAVIFLLRREKNTLRLLDDKKFIFNLMVCGRMNDN 1081

Qy        918 LS------GSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLE---HLMDIF 968
              ||       : : | : |  |       :| :||              |:: |:   |  ::
Db       1082 LSLEELVLHTSAPLDTAVRLSRALTLCALREKE-----------RSVDLQAAAHHCELM 1129

Qy        969 EREQITLD-----------VLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNP 1016
                : ::|             :|  : |: |  :             ||:| :: :  |  :
Db       1130 ASDLLSLSASAGGQGAGAGPILRALDHRGLSVLDC--------LIEGRQKGVVSQPAVQT 1181

Qy       1017 YLT 1019
              |||
Db       1182 YLT 1184


RESULT 36
Q9UFA4
ID   Q9UFA4      PRELIMINARY;      PRT;  791 AA.
AC   Q9UFA4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP434D2328.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
RA   Wiemann S.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL133087; CAB61404.1; -.
DR   PIR; T42691; T42691.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 22.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 19.
DR   PROSITE; PS50088; ANK_REPEAT; 17.
```

```
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
FT   NON_TER       1       1
SQ   SEQUENCE   791 AA;  85632 MW;  43626C2F4DF558EF CRC64;

  Query Match             10.2%;  Score 657;  DB 4;  Length 791;
  Best Local Similarity   28.2%;  Pred. No. 6.3e-33;
  Matches  243;  Conservative  115;  Mismatches  331;  Indels  172;  Gaps   29;

Qy        102 LFEACRNGDVERVKRLVT-PEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
              |  |  || :  || |:    :::  :  |  :| || |   |:  ||  |:  ||||
Db          7 LHAAASNGQINVVKHLLNLGVEIDEINVYG--NTALHIACYNGQDAVVNELIDYGANVNQ 64

Qy        161 RDDGGLIPLHNACSFGH-AEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQ 219
              :: |  ||| | :  | |  : ||: :||| | :     :|||  |: |:       |:|
Db         65 PNNNGFTPLHFAAASTHGALCLELLVNNGADVNIQSKDGKSPLHMTAVHGRFTRSQTLIQ 124

Qy        220 HGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT------- 272
              :|  |    : || | | :|                  || |:|  :  |:|
Db        125 NGGEIDCVDKDGNTPLHVA-------------------ARYGHELLINTLITSGADTAK 164

Qy        273 -------PLN---VNCHA-------SDGRK-------------------STP-------L 289
                    ||:   :| |:        | |:|                   ||         |
Db        165 CGIHSMFPLHLAALNAHSDCCRKLLSSGQKYSIVSLFSNEHVLSAGFEIDTPDKFGRTCL 224

Qy        290 HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDL 349
              | ||     |: ::||   ||| | |||    ||| | :  |:    | ||  || ||  |
Db        225 HAAAAGGNVECIKLLQSSGADFHKKDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDD 284

Qy        350 WQFTPLHEAASKNRVEVCSLLLSYGADPTLL-NCHNKSAIDLAPTPQLKER---LAYEFK 405
              |  | || ||: :               : |:| | |: |  :|    :|||:   |  ||
Db        285 WGRTALHYAAASD----------MDRNKTILGNAHDNSE-ELERARELKEKEATLCLEF- 332

Qy        406 GHSLLQAARFADVTRIKKHLSLEMVNFKHPQTHE----TALHCAAASPYPKRKQICELLL 461
              |||                  |  :|   :     ::| |||  :   :|  ||||
Db        333 ---LLQ-----------------NDANPSIRDKEGYNSIHYAAAYGH---RQCLELLL 367

Qy        462 RKGANINEKTKEFLT--PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGH 519
              :  :  |::     |  |||:|:   |:  :||:::    ::  |  |:|:|  ||: ||
Db        368 ERTNSGFEESDSGATKSPLHLAAYNGHHQALEVLLQSLVDLDIRDEKGRTALDLAAFKGH 427

Qy        520 LQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV 579
              :    |::           || :     |   :  |   :  |::   | ||| |
Db        428 TECVEALIN----------QGASIFVKDNVTKRTPLHASVINGHTLCLRLLLEIAD---- 473

Qy        580 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
                    ::|: :| :|:  ||| |  |  :  |  ||:  |:|   |  |  ||
Db        474 -------NPEAVDVKDAKGQ--TPLMLAVAYGHIDAVSLLLEKEANVDTVDILGCTALHR 524

Qy        640 ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHG---ADPTKKN 696
              || |  ::|::    :   |    |||| |||:|       |||       |   |:
Db        525 GIMTGHEECVQMLLEQEVSILCKDSRGRTPLHYAAARGHATWLSELLQMALSEEDCCFKD 584

Qy        697 RDGNTPLD--LVKDGDTDIQDLLR--------GD-------AALLDAAKKGCLARVKKLS 739
              | |||          :  |: ||          |:       | : |      | : :
Db        585 NQGYTPLHWACYNGNENCIEVLLEQKCFRKFIGNPFTPLHCAIINDHG--NCASLLLGAI 642

Qy        740 SPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHV 799
              |:||| :||  |||| ||  :::|  : ||:| | ||| |   |    |  || |
Db        643 DSSIVSCRDDKGR--TPLHAAAFADHVECLQLLLRHSAPVNAVDNSGKTALMMAAENGQA 700

Qy        800 DVAALLI-KYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTL---KNQEGQTP 855
                :|:    | :   ||   |||| |  ||  :   |:|    | :|   ||  |||
Db        701 GAVDILVNSAQADLTVKDKDLNTPLHLACSKGHEKCALLILDKIQDESLINEKNNALQTP 760

Qy        856 LDLVSADDVSALLTAAMPPSA 876
              | : : : :  ::   :  |
Db        761 LHVAARNGLKVVVEELLAKGA 781


RESULT 37
Q9VMR4
ID   Q9VMR4      PRELIMINARY;      PRT;  1619 AA.
AC   Q9VMR4; Q9NGR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG11020 protein (Mechanosensory transduction channel NOMPC).
GN   NOMPC OR CG11020 OR CG17468.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
```

```
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Walker R.G., Willingham A.T., Zuker C.S.;
RT   "A Drosophila mechanosensory transduction channel.";
RL   Science 0:0-0(2000).
DR   EMBL; AE003609; AAF52248.3; -.
DR   EMBL; AF242296; AAF59842.1; -.
DR   HSSP; Q00420; 1AWC.
DR   FlyBase; FBgn0016920; nompC.
DR   GO; GO:0007605; P:hearing; IMP.
DR   GO; GO:0007638; P:mechanosensory behavior; IMP.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002111; Cat_channel_TrpL.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR000550; Hppk.
DR   Pfam; PF00023; ank; 29.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 28.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
DR   PROSITE; PS00794; HPPK; 1.
KW   ANK repeat; Ionic channel; Repeat; Transmembrane.
SQ   SEQUENCE   1619 AA;  176844 MW;  B4FE239ABB1D1138 CRC64;

  Query Match             9.8%;  Score 635.5;  DB 5;  Length 1619;
  Best Local Similarity   25.6%;  Pred. No. 4.5e-31;
  Matches  260;  Conservative  134;  Mismatches  340;  Indels  281;  Gaps   33;

Qy         29 DPVTAG---SQAARALSASSPGG-----LALLLAGPGLLLRLLA-------LLLAVAAAR 73
Db        186 DPFSTGGSRSQTAVHLVSSRQTGTATNILRALLAAAGKDIRLKADGRGKIPLLLAVE--- 242

Qy         74 IMSGRRCAGGGAAC----ASAAAEAVEPAARE----LPEACRNGDVERVKRLVTPEKVNS 125
Db        243 -------SGNQSMCRELLAAQTAEQLKATTANGDTALHLAARRRDVDMV-RILVDYGTNV 294

Qy        126 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLL- 184
Db        295 DTQNGEGQTPLHIAAAEGDEALLKYFYGVRASASIADNQDRTPMHLAAENGHAHVIEILA 354

Qy        185 ----------------LRHGADPNA----------RDNWNYTP-------LHEAAIKGKI 211
Db        355 DKFKASIFERTKDGSTLMHIASLNGHAECATMLFKKGVYLHMPNKDGARSIHTAAAYGHT 414

Qy        212 DVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLT--------------GEYKKDELLE 257
Db        415 GIINTLLQKGEKVDVTTNDNYTALHIAVESAKPAVVETLLGFGADVHVRGGKLRETPLHI 474

Qy        258 SAR--SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY-NRVKIVQLLLQHGADVHAK 314
Db        475 AARVKDGDRCALMLLKSGASPNLTTDD--CLTPVHVAARHGNLATLMQLLEDEGDPLYKS 532

Qy        315 DKGDLVPLHNACSYGHYEVTELLV-----KHG-----ACVNAMDLWQFTPLHEAASKNRV 364
Db        533 NTGE-TPLHMACRACHPDIVRHLIETVKEKHGPDKATTYINSVNEDGATALHYTCQITKE 591

Qy        365 EV---------CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 415
Db        592 EVKIPESDKQIVRMLLENGADVTL---QTKTALETA--------------FHYCAVAGN 633

Qy        416 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 475
Db        634 NDVL-------MEMISHMNPTDIQKAM------------------------NRQSSVGW 661

Qy        476 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 535
Db        662 TPLLIACHRGHMELVNNLLANHARVDVFDTEGRSALHLAAERGYLHVCDALLTNKAPINS 721

Qy        536 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 595
Db        722 KSRVGRTALHL------------------------AAMNGFTHLVKFLIKDHNAVIDI 755

Qy        596 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL-LVK 654
Db        756 LTLRKQTPLHLAAASGQMEVCQLLLELGANIDATDDLGQKPIHVAAQNNYSEVAKLFLQQ 815

Qy        655 HGAVVN--------------------VADLWKF---------------TPLHEAAAKGK 678
Db        816 HPSLVNATSKDGNTCAHIAAMQGSVKVIEELMKFDRSGVISARNKLTDATPLQLAAEGGH 875

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKK-------- 729
Db        876 ADVVKALVRAGASCTEENKAGFTAVHLAAQNGHGQVLDVLKSTNSLRINSKKLGLTPLHV 935

Qy        730 ----GCLARVKKL--SSPDNVNCRDTQGRH----------STPLHLAA------------ 761
Db        936 AAYYGQADTVRELLTSVPATVKSETPTGQSLFGDLGTESGMTPLHLAAFSGNENVVRLLL 995

Qy        762 -------------GYNNLEVAEY---------LLQHGAD-VNAQDKGGLIPLHNAASYG 797
Db        996 NSAGVQVDAATIENGYNPLHLACFGGHMSVVGLLLSRSAELLQSQDRNGRTGLHIAAMHG 1055

Qy        798 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 852
Db       1056 HIQMVEILLGQGAEINATDRNGWTPLHCAAKAGHLEVVKLLCEAGASPKSETNYG 1110

RESULT 38
Q13484
ID   Q13484      PRELIMINARY;      PRT;  1088 AA.
AC   Q13484;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ankyrin G119.
GN   ANK3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=96234670; PubMed=8666667;
RA   Devarajan P., Stabach P.R., Mann A.S., Ardito T., Kashgarian M.,
RA   Morrow J.S.;
RT   "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney
RT   and muscle that binds beta I sigma spectrin and associates with the
RT   Golgi apparatus.";
RL   J. Cell Biol. 133:819-830(1996).
DR   EMBL; U43965; AAB08437.1; -.
DR   HSSP; P55273; 1BI8.
DR   GO; GO:0005783; C:endoplasmic reticulum; TAS.
DR   GO; GO:0005794; C:Golgi apparatus; TAS.
DR   GO; GO:0007016; P:cytoskeletal anchoring; TAS.
DR   GO; GO:0006899; P:nonselective vesicle transport; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 14.
```

```
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 12.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 12.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1088 AA;  119428 MW;  769C88D40A78DE86 CRC64;

  Query Match             9.7%;  Score 629.5;  DB 4;  Length 1088;
  Best Local Similarity   28.1%;  Pred. No. 5.8e-31;
  Matches  195;  Conservative   85;  Mismatches  234;  Indels  181;  Gaps   16;

Qy        304 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 363
              ||   |: :||       ||| ||     :| |||:|||| : |:     ||:| ||
Db          7 LLDKKANPNAKALNGFTPLHIACKKNRIKVMELLLKHGASIQAVTESGLTPIHVAAFMGH 66

Qy        364 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 423
              | : | |: :|| |   |
Db         67 VNIVSQLMHHGASPNTTNVRG----------------------------------------- 87

Qy        424 HLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASE 483
                          ||||| || |    : ::   |:: || :  | |:  ||||:::
Db         88 --------------ETALHMAARS---GQAEVVRYLVQDGAQVEAKAKDDQTPLHISAR 129

Qy        484 KAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTA 543
                 |:|: :::  |  ||   | | || :|  ||      || :|   :| : :||
Db        130 LGKADIVQQLLQQGASPNAATTSGYTPLHLSAREGHEDVAAFLLDHGASLSITTKKGF-- 187

Qy        544 LQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTP 603
                                                                    ||
Db        188 ------------------------------------------------------TP 189

Qy        604 LHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVAD 663
              || || | :: |   |||  |   |  | || ||| |  | : :|| ||:  ||  : |
Db        190 LHVAAKYGKLEVANLLLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAA 249

Qy        664 LWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAA 722
                :|||| || | : :|   ||:|||      | |   : |  ::|  |:  || |  |
Db        250 KNGYTPLHIAAKKNQMDIATTLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNA 309

Qy        723 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 782
                          ||| : |  |||||||  : : ||| |:  || |:||
Db        310 -------------------NVNLSNKSG--LTPLHLAAQEDRVNVAEVLVNQGAHVDAQ 347

Qy        783 DKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHG 842
              | |  |||     ||:: :   |::::| |||   |  :||||:|||:| | :  :|| :
Db        348 TKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTPLHQAAQQGHTHIINVLLQNN 407

Qy        843 ADPTLKNQEGQTPLDL------VSADDVSALLT-AAMPPSALPSCYK---PQVLNGV--R 890
              | |     | | | :      :|  |   ::|   |  : :   :|   |: :| | |
Db        408 ASPNELTVNGNTALGIARRLGYISVVDTLKIVTEETMTTTTVTEKHKMNVPETMNEVLDM 467

Qy        891 SPGATADALSS------GP---------SSPS------SLSAASSLDNLSGSFSELSSVV 929
              |      ||::       ||          | |:      || | |: |    |:    ::
Db        468 SDDEGEDAMTGDTDKYLGPQDLKELGDDSLPAEGYMGFSLGARSASDR---SY----TLN 520

Qy        930 SSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 964
                ||    :  :|:  ||  :  :| | |    : |
Db        521 RSSYARDSMMIEELLVPSKEQHLT-FTREFDSDSL 554


RESULT 39
Q9GRV5
ID   Q9GRV5      PRELIMINARY;      PRT;  1913 AA.
AC   Q9GRV5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Y71A12B.4 protein.
GN   Y71A12B.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sulston J.E.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; AL132902; CAC14420.1; -.
DR   HSSP; Q00420; 1AWC.
DR   WormPep; Y71A12B.4; CE26238.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005262; F:calcium channel activity; IEA.
DR   GO; GO:0006816; P:calcium ion transport; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002111; Cat_channel_TrpL.
DR   InterPro; IPR002153; Trans_receptor.
DR   Pfam; PF00023; ank; 29.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR01097; TRNSRECEPTRP.
DR   SMART; SM00248; ANK; 24.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Ionic channel; Repeat; Transmembrane.
SQ   SEQUENCE   1913 AA;  211430 MW;  D421B7F253C42A27 CRC64;

  Query Match             9.7%;  Score 627;  DB 5;  Length 1913;
  Best Local Similarity   24.5%;  Pred. No. 2e-30;
  Matches  289;  Conservative  154;  Mismatches  450;  Indels  288;  Gaps   38;

Qy         15 QRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARI 74
              ::||||      | |  : |  ::|   | | |   :| |   ||
Db        327 RKGARV--------DYSTKDGRNATHIAAMYSGVETLEL----ILKRYSELL-------- 366

Qy         75 MSGRRCAGG------GAACASAAAEAVEPAARELFE-----------------ACRNGD 110
                |: ||        ||   : :|  |  : : |                  | : |:
Db        367 ---RKGAGPKKQLAIHVACERKSKKAF-PIVKRILEDTDQRMAEDGDGSLPIHLAFKFGN 422

Qy        111 VERVKRLVT-PEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNG---ANVQARDDGGL 166
              |  |: |:: |    :|    |   | || ||   |  : |   :  |   |||| |   |
Db        423 VNIVELLLSGPSDEQTRKADGNGDTLLHLAARSGNIEAVRTAIAAGCDNANVQNR--VGR 480

Qy        167 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLL-QHGAEPT 225
              |||      |   :: :: :  || |  |  : ||:| || :|   :   |: : |
Db        481 TPLHEVAEVGDQNMLKIMFKLRADANIHDKEDKTPVHVAAERGDTSMVESLIDKFGGSIR 540

Qy        226 IRNTDGRTALDLADPSAKAVLTGEY----------KKDEL-LESARSGNEEKMMALLTP 273
              |  || | | :|  |        :            ||  | | || :     :: :|
Db        541 ARTRDGSTLLHIAACSGHTSTALAFLKRGVPLFMPNKKGALGLHSAAAAGFNDVVKMLIA 600

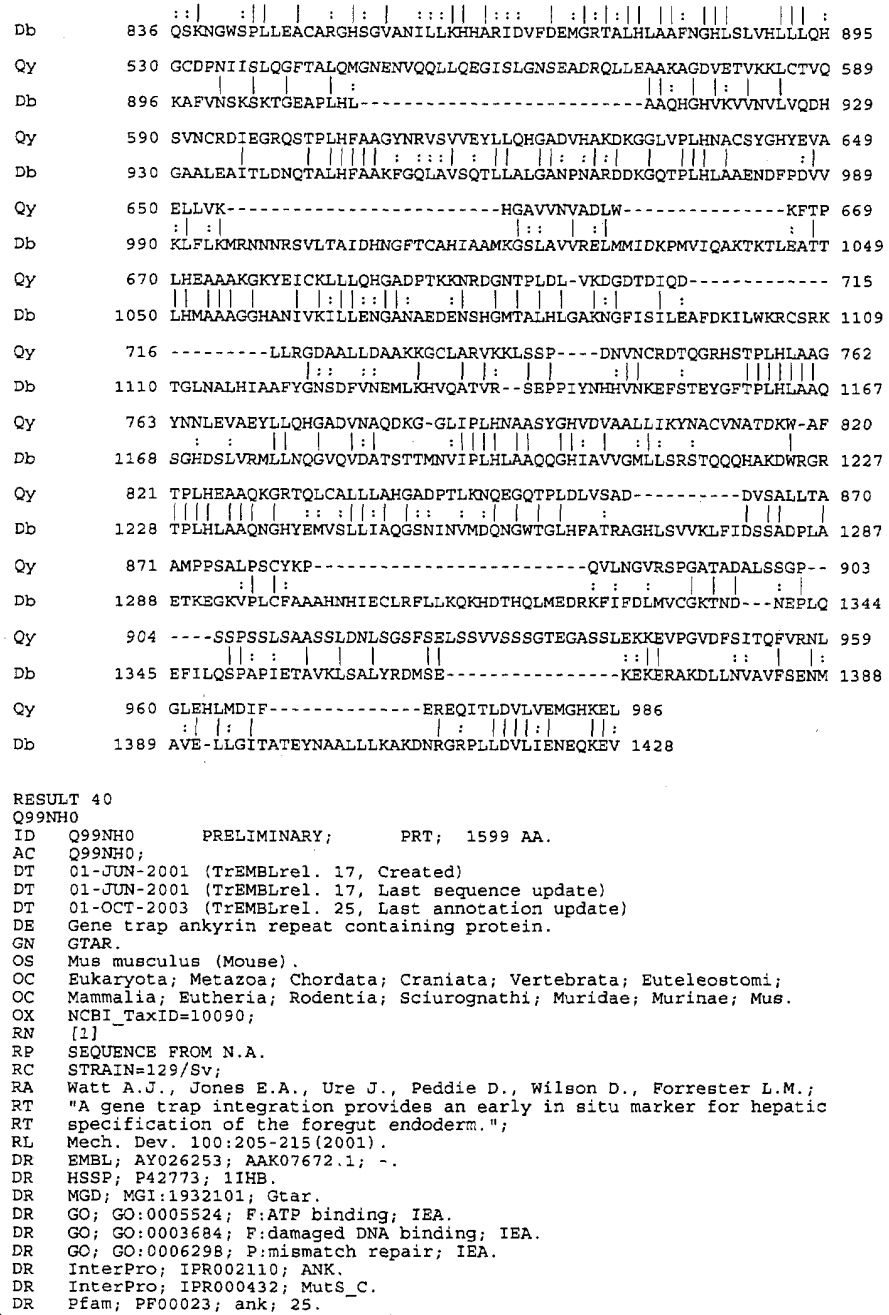
Qy        274 LNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDL--VPLHNACSYGHY 331
                |          | ||:|    :  :|: ||  |||:| |  |:|     || | |
Db        601 RGTNVDVRTRDNYTALHVAVQSGKASVVETLLGSGADIHVKG-GELGQTALHIAASLNGA 659

Qy        332 EVTE---LLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTL--------- 379
              |  :   :|:| |   :   :  | || ||     ::  |||   ||  :
Db        660 ESRDCAMMLLKSGGQPDVAQMDGETCLHIAARSGNKDIMRLLLDENADSKISSKIGETPL 719

Qy        380 ----LNCHNKSA-------IDLAPTPQLKERLAYEFK-GHSLLQAAREADVTRIKKHLSL 427
                  :|: ::|         ::    |||| : :    | : |  |  |:: : : |
Db        720 QVAAKSCNFEAASMILKHLSEVLTQEQLKEHVNHRTNDGFTALHYA--AEIEQRQLHFPG 777

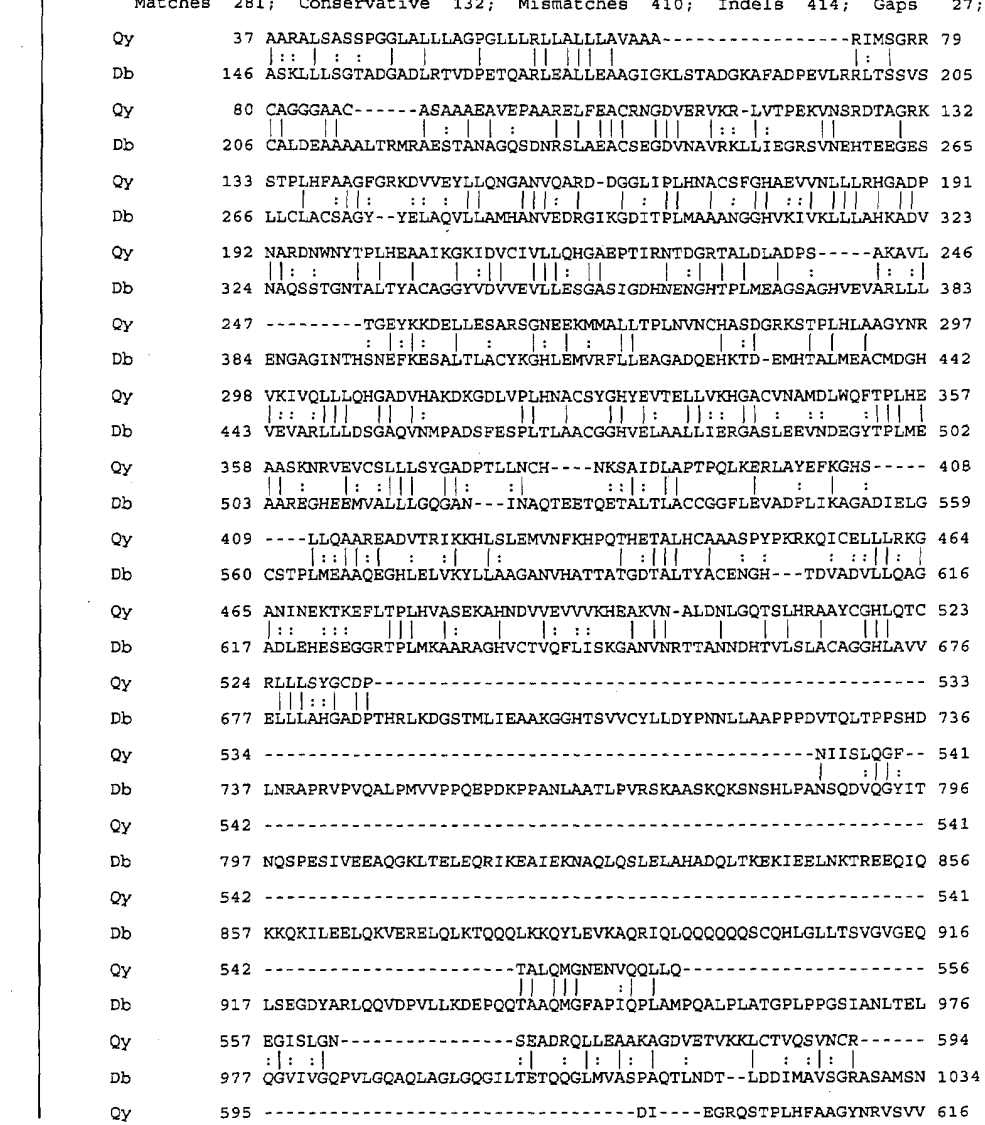
Qy        428 E-------------MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI-----NE 469
              |              ||       :|||:| || |    :  :  :: : ||       |:
Db        778 EDAKLVNLLIDYGGMVEMPSLNANETAMHMAARS--GNQAVLLAMVNKIGAGAVQIVQNK 835

Qy        470 KTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSY 529
```



RESULT 40
Q99NH0
ID Q99NH0 PRELIMINARY; PRT; 1599 AA.
AC Q99NH0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Gene trap ankyrin repeat containing protein.
GN GTAR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Watt A.J., Jones E.A., Ure J., Peddie D., Wilson D., Forrester L.M.;
RT "A gene trap integration provides an early in situ marker for hepatic
RT specification of the foregut endoderm.";
RL Mech. Dev. 100:205-215(2001).
DR EMBL; AY026253; AAK07672.1; -.
DR HSSP; P42773; 1IHB.
DR MGD; MGI:1932101; Gtar.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0006298; P:mismatch repair; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000432; MutS_C.
DR Pfam; PF00023; ank; 25.
DR Pfam; PF00488; MutS_V; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 25.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
SQ SEQUENCE 1599 AA; 170271 MW; CCBB6C56D8ECB7C8 CRC64;

Query Match 9.1%; Score 590; DB 11; Length 1599;
Best Local Similarity 22.7%; Pred. No. 3.4e-28;
Matches 281; Conservative 132; Mismatches 410; Indels 414; Gaps 27;



```
Db       1035 TPTHSIAASVSQPQTPTPSPIISPSAMLPIYPAIDIDAQTESNHDTALTLACAGGHEELV 1094

Qy        617 EYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNV-ADLWKFTPLHEAAA 675

Db       1095 QTLLERGASIEHRDKKGFTPLILAATAGHVGVVEILLDNGADIEAQSERTKDTPLSLACS 1154

Qy        676 KGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD--TDIQDLLR------------GD 720

Db       1155 GGRQEVVELLLARGANKEHRNVSDYTPLSLAASGGYVNIIKILLNAGAEINSRTGSKLGI 1214

Qy        721 AALLDAAKKGCLARVK-KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 779

Db       1215 SPLMLAAMNGHTAAVKLLLDMGSDINAQIETNR-NTALTLACFQGRTEVVSLLLDRKANV 1273

Qy        780 NAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA--TDKWAFTPLHEAAQKGRTQLCAL 837

Db       1274 EHRAKTGLTPLMEAASGGYAEVGRVLLDKGADVNAPPVPSSRDTALTIAADKGHYKFCEL 1333

Qy        838 LLAHGADPTLKNQEGQTPLDLVSAD---DVSALLTAA 871

Db       1334 LIGKGAHIDVRNKKGNTPLWLAANGGHLDVVQLLVQA 1370


RESULT 41
O75179
ID   O75179      PRELIMINARY;      PRT;  2486 AA.
AC   O75179;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein KIAA0697 (Fragment).
GN   KIAA0697.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA   Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
CC   -!- SIMILARITY: CONTAINS 1 KH DOMAIN.
DR   EMBL; AB014597; BAA31672.2; -.
DR   PIR; T00353; T00353.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003684; F:damaged DNA binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0006298; P:mismatch repair; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000432; MutS_C.
DR   Pfam; PF00023; ank; 25.
DR   Pfam; PF00013; KH; 1.
DR   Pfam; PF00488; MutS_V; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 25.
DR   SMART; SM00322; KH; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50084; KH_TYPE_1; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
FT   NON_TER       1      1
SQ   SEQUENCE   2486 AA;  263244 MW;  217C5B039240EEA9 CRC64;

  Query Match             9.1%;  Score 588.5;  DB 4;  Length 2486;
  Best Local Similarity   22.5%;  Pred. No. 8.5e-28;

  Matches  286;  Conservative  136;  Mismatches  415;  Indels  435;  Gaps   29;

Qy         37 AARALSASSPGGLALLLAGPGLLLRLLALLLAVAAA----------------RIMSGRR 79

Db         34 ASKLLLSGTADGADLRTVDPETQARLEALLEAAGIGKLSTADGKAFADPEVLRRLTSSVS 93

Qy         80 CAGGGAAC------ASAAAEAVEPAARELFEACRNGDVERVKR-LVTPEKVNSRDTAGRK 132

Db         94 CALDEAAAALTRMRAESTANAGQSDNRSLAEACSEGDVNAVRKLLIEGRSVNEHTEEGES 153

Qy        133 STPLHFAAGFGRKDVVEYLLQNGANVQARD-DGGLIPLHNACSFGHAEVVNLLLRHGADP 191

Db        154 LLCLACSAGY--YELAQVLLAMHANVEDRGIKGDITSLMAAANGGHVKIVKLLLAHKADV 211

Qy        192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPS-----AKAVL 246

Db        212 NAQSSTGNTALTYACAGGYVDVVKVLLESGASIEDHNENGHTPLMEAGSAGHVEVARLLL 271

Qy        247 ---------TGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNR 297

Db        272 ENGAGINTHSNEFKESALTLACYKGHLEMVRFLLEAGADQEHKTD-EMHTALMEACMDGH 330

Qy        298 VKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHE 357

Db        331 VEVARLLLDSGAQVNMPADSFESPLTLAACGGHVELAALLIERGASLEEVNDEGYTPLME 390

Qy        358 AASKNRVEVCSLLLSYGADPTLLNCH----NKSAIDLAPTPQLKERLAYEFKGHS----- 408

Db        391 AAREGHEEMVALLLGQGAN---INAQTEETQETALTLACCGGFLEVADFLIKAGADIELG 447

Qy        409 ----LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKG 464

Db        448 CSTPLMEAAQEGHLELVKYLLAAGANVHATTATGDTALTYACENGH---TDVADVLLQAG 504

Qy        465 ANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN-ALDNLGQTSLHRAAYCGHLQTC 523

Db        505 ADLEHESEGGRTPLMKAARAGHVCTVQFLISKGANVNRTTANNDHTVLSLACAGGHLAVV 564

Qy        524 RLLLSYGCDP------------------------------NIIS--------------- 537

Db        565 ELLLAHGADPTHRLKDGSTMLIEAAKGGHTSVVCYLLDYPNNLLSAPPPDVTQLTPPSHD 624

Qy        538 ------------------------------------------------------LQGF-- 541

Db        625 LNRAPRVPVQALPMVVPPQEPDKPPANVATTLPIRNKAASKQKSSSHLPANSQDVQGYIT 684

Qy        542 ------------------------------------------------------------ 541

Db        685 NQSPESIVEEAQGKLTELEQRIKEAIEKNAQLQSLELAHADQLTKEKIEELNKTREEQIQ 744

Qy        542 ------------------------------------------------------------ 541

Db        745 KKQKILEELQKVERELQLKTQQQLKKQYLEVKAQRIQLQQQQQQSCQHLGLLTPVGVGEQ 804

Qy        542 ----------------------TALQMGNENVQQLLQ---------------------- 556

Db        805 LSEGDYARLQQVDPVLLKDEPQQTAAQMGFAPIQPLAMPQALPLAAGPLPPGSIANLTEL 864

Qy        557 EGISLGN----------------SEADRQLLEAAKAGDVETVKKLCTVQSVNCR------ 594

Db        865 QGVIVGQPVLGQAQLAGLGQGILTETQQGLMVASPAQTLNDT--LDDIMAVSGRASAMSN 922

Qy        595 ---------------------------------DI----EGRQSTPLHFAAGYNRVSVV 616

Db        923 TPTHSIAASISQPQTPTPSPIISPSAMLPIYPAIDIDAQTESNHDTALTLACAGGHEELV 982

Qy        617 EYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNV-ADLWKFTPLHEAAA 675

Db        983 QTLLERGASIEHRDKKGFTPLILAATAGHVGVVEILLDNGADIEAQSERTKDTPLSLACS 1042

Qy        676 KGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD--TDIQDLLR------------GD 720

Db       1043 GGRQEVVELLLARGANKEHRNVSDYTPLSLAASGGYVNIIKILLNAGAEINSRTGSKLGI 1102
```

```
Qy        721 AALLDAAKKGCLARVK-KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 779
Db       1103 SPLMLAAMNGHTAAVKLLLDMGSDINAQIETNR-NTALTLACFQGRTEVVSLLLDRKANV 1161

Qy        780 NAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA--TDKWAFTPLHEAAQKGRTQLCAL 837
Db       1162 EHRAKTGLTPLMEAASGGYAEVGRVLLDKGADVNAPPVPSSRDTALTIAADKGHYKFCEL 1221

Qy        838 LLAHGADPTLKNQEGQTPLDLVSAD---DVSALLTAAMPPSALPSCYKPQVLNGVRSPGA 894
Db       1222 LIGRGAHIDVRNKKGNTPLWLAANGGHLDVVQLLVQA--------------------GA 1260

Qy        895 TADALSSGPSSP 906
Db       1261 DVDAADNRKITP 1272
```

RESULT 42
O83807
ID O83807 PRELIMINARY; PRT; 934 AA.
AC O83807;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ankyrin, putative.
GN TP0835.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
DR EMBL; AE001254; AAC65803.1; -.
DR PIR; H71274; H71274.
DR HSSP; P25963; 1IKN.
DR TIGR; TP0835; -.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00023; ank; 23.
DR SMART; SM00248; ANK; 14.
DR PROSITE; PS50088; ANK_REPEAT; 15.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW Complete proteome.
SQ SEQUENCE 934 AA; 100687 MW; 098E5F9F0165A2BB CRC64;

Query Match 8.7%; Score 563.5; DB 16; Length 934;
Best Local Similarity 23.8%; Pred. No. 7e-27;
Matches 254; Conservative 159; Mismatches 373; Indels 281; Gaps 36;

```
Qy         86 ACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRK 145
Db         24 SCTTTSSPSPQ-ASVDLFEMVARGDLDTLKTRFGKDSVNVRDANG--DTLLH--VGVLRN 78

Qy        146 D--VVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLH 203
Db         79 DAHVVDFLLSMGADTEAQDASGSTPLMVAVTENCFESVRMLIARDASLFSRDAQGDTPLT 138

Qy        204 EAAIKGKIDVCIV------------------------LLQH----GAEPTIRNTDGRTA 234
```

```
Db        139 RAIRKGTAHELVTRKTLLQKDKSGKTPWHWAVRALDRDLIKHLVTLGPPTQERDADGHTP 198

Qy        235 LDLADPSAKAVLTGEYKKDELLESAR--SGNEEKMMALLTPLNVNCHASDGRKSTPLHLA 292
Db        199 LSLAYSSSSDARAAEVAADLLLGGAELMHGSFSEFEIAVLKRNYDLHFTDG--STPLHVM 256

Qy        293 AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQF 352
Db        257 AKRGYTGFVQFLVDRKVNLNAK--------------------------------NLSSA 283

Qy        353 TPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL-APTPQLKERLAYEFKGHSLLQ 411
Db        284 TPLHEAVRAGQVDAAVLLLRSGADPNVRDASGNTCLHLVAPAP-FRVRLV-----GALLD 337

Qy        412 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 471
Db        338 AGASVAI--------------KDDYGETPLHVAAR--LGMDRAFVERLVGAGADISERN 380

Qy        472 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS--- 528
Db        381 KKGETPLVLTIDRDHRDLTAYFVSLGADIHAEDMRGETPLTKALARG-LETVKIVVTDSN 439

Qy        529 ------YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV 582
Db        440 LYKQDVVGRDPLHVAVSRRAHLDI----VKFLFRE---------PKQMI---------- 475

Qy        583 KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGAD---------------- 625
Db        476 ----------ARDTMG--NTLLHYAVANDDRAVGEFLMREGADIFSTNVHGVSPLKTALT 523

Qy        626 --------------VHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 671
Db        524 TSGGREDWILTAANVHAQDTGGNTPLHLACEWKLTQAINGILRKGAEIEARNLNQETPLF 583

Qy        672 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 731
Db        584 SAVKSDAAEVISILLHPQA-------------------------GNPALVDA----- 610

Qy        732 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLL-----QHGADVNAQDKGG 786
Db        611 --------------RDAVG--NTVLHACVRWSALRSADVLIREADARHVSLLNARNLSG 653

Qy        787 LIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPT 846
Db        654 KPPLHLAARAGNVDFIRLLLSHRVALHMGDETGKSALTDAVLADQEESVHMLLSAGANPV 713

Qy        847 LKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSP 906
Db        714 QQDMYGRTPLH-------EAVL-----------CNSQSVIAALRAAGGNPFARDSYGTTP 755

Qy        907 SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLE---KKEVPGVDFSITQFVRNLGLEH 963
Db        756 LSLALLKG-DTFVGAVVGKDPLLANS--DGQTPLHLAVMENVPVQTF-------RLLLAK 805

Qy        964 LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIK-GVERLISGQQGLNP---YLT 1019
Db        806 GYPIDKRDRMGSSALVLAIKKDRSDL-------CHELLALGADLFIANNVGESPALLVLS 858

Qy       1020 LNTSGSGTIL-IDLSPDDKEFQSV---------EEEMQSTVREHRDG 1056
Db        859 KNTSILKTLVGFAVNKTDSAGESILHYAAKVADEKTLQGLLAMNRFG 905
```

RESULT 43
Q8C036
ID Q8C036 PRELIMINARY; PRT; 531 AA.
AC Q8C036;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical ankyrin-repeat containing protein.
GN G431002C21RIK.

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK032429; BAC27865.1; -.
DR   PIR; PT0649; PT0712.
DR   PIR; PT0706; PT0706.
DR   MGD; MGI:2444029; G431002C21Rik.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 15.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 14.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   531 AA;  57007 MW;  8CF4FFA30CDF48AC CRC64;

  Query Match             8.6%;  Score 558.5;  DB 11;  Length 531;
  Best Local Similarity   30.2%;  Pred. No. 6e-27;
  Matches  178;  Conservative   78;  Mismatches  224;  Indels  109;  Gaps   14;

Qy        102 LFEACRNGDVERVKRLVT-PEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
Db         12 LVQAIFSRDVEEVRSLLSQKENINVLDQERR--TPLHAAAYVGDVPILQLLLMSGANVNA 69

Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQH 220
Db         70 KDTLWLTPLHRAAASRNEKVLGLLLAHSADVNARDKLWQTPLHVAA-------------- 115

Qy        221 GAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHA 280
Db        116 ----------------ANRATKCA------------------EALAPLLSSLNV---- 135

Qy        281 SDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKH 340
Db        136 ADRSGRSALHHAVHSGHLETVNLLLNKGASLNVCDKKERQPLHWAAFLGHLEVLKLLVAR 195

Qy        341 GACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERL 400
Db        196 GADLSCKDRKGYGLLHTAAASGQIEVVKHLLRMGAEIDEPNAFGNTAL------------ 243

Qy        401 AYEFKGHSLLQAAREADVTRIKKHLSLEMV----NFKHPQTHE-TALHCAAASPYPKRKQ 455
Db        244 ------HIACYLGQDA--------VAIELVNAGANVNQPNDKGFTPLHVAAVS---TNGA 286

Qy        456 IC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRA 514
Db        287 LCLELLVNNGADVNYQSKEGKSPLHMAAIHGRFTRSQILIQNGSEIDCADKFGNTPLHVA 346

Qy        515 AYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAA 574
Db        347 ARYGHELLISTLMTNGADTARRGIHDMFPLHL-----------AVLFGFSDCCRKLLSSG 395

Qy        575 K----AGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKD 630
Db        396 QLYSIVSSLSNEHVLSAGLDINTPDSLGR--TCLHAAASGGNVECLNLLLSSGADLRRRD 453

Qy        631 KGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 679
Db        454 KFGRTPLHYAAANGSYQCAVTLVTAGAGVNEADCKGCSPLHYAAASDTY 502


RESULT 44
Q8IWZ3
```

```
ID   Q8IWZ3      PRELIMINARY;      PRT;  2542 AA.
AC   Q8IWZ3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Multiple ankyrin repeats single KH domain protein isoform 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Poulin F., Sonenberg N.;
RT   "Two overlapping reading frames in the second exon of the
RT   translational inhibitor 4E-BP3.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF521882; AAO14943.1; -.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00023; ank; 25.
DR   Pfam; PF00013; KH; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00322; KH; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50084; KH_TYPE_1; 1.
SQ   SEQUENCE   2542 AA;  269456 MW;  AB310E826A4134D0 CRC64;

  Query Match             8.6%;  Score 558.5;  DB 4;  Length 2542;
  Best Local Similarity   21.8%;  Pred. No. 7.1e-26;
  Matches  293;  Conservative  122;  Mismatches  418;  Indels  513;  Gaps   33;

Qy          5 RRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALS----ASSPGGLALLLAGPGLLL 60
Db         85 RTGGGGGAGSGSDEDEVSEVESFILDQEDLDNPVLKTTSEIFLSSTAEGADLRTVDPETQA 144

Qy         61 RLLALLLAVAAA----------------RIMSGRRCAGGGAACASAAAEAV------EP 97
Db        145 RLEALLEAAGIGKLSTADGKAFADPEVLRRLTSSVSCALDEAAAALTRMKAENSHNAGQV 204

Qy         98 AARELFEACRNGDVERVKRLVTP-EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGA 156
Db        205 DTRSLAEACSDGDVNAVRKLLDEGRSVNEHTEEGESLLCLACSAGY--YELAQVLLAMHA 262

Qy        157 NVQAR-DDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCI 215
Db        263 NVEDRGNKGDITPLMAASSGGYLDIVKLLLLHDADVNSQSATGNTALTYACAGGFVDIVK 322

Qy        216 VLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT-PL 274
Db        323 VLLNEGANIEDHNENGHT-------------------PLMEAASAGHVEVARVLLDHGA 362

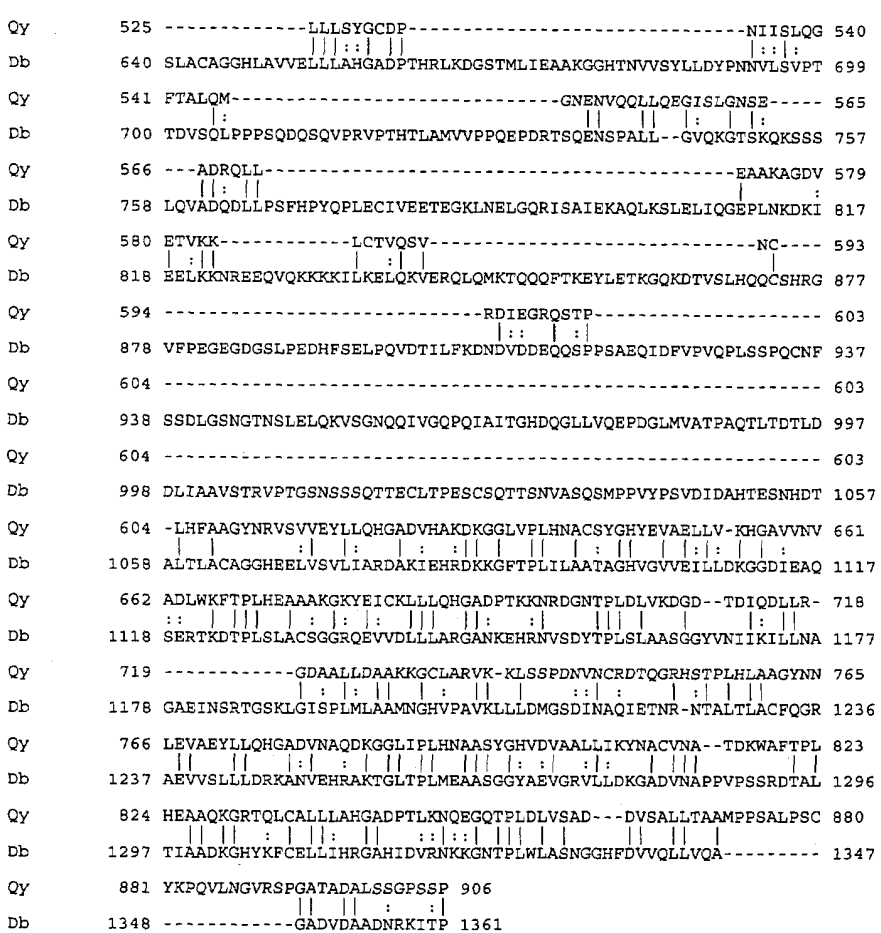
Qy        275 NVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVT 334
Db        363 GINTHSNEFKESA-LTLACYKGHLDMVRFLLEAGADQEHKTDEMHTALMEACMDGHVEVA 421

Qy        335 ELLVKHGACVNAMDLWQF-TPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPT 393
Db        422 RLLLDSGAQVN-MPADSFESPLTLAACGGHVELAALLIERGANLEEVN-------DEGYT 473

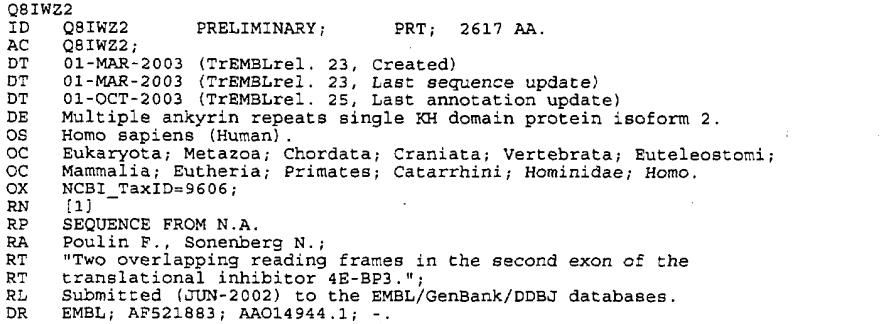
Qy        394 PQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETALHCAAASPYPK 452
Db        474 P--------------LMEAAREGHEEMVALLLAQGANINAQTEETQETALTLACCGGFSE 519

Qy        453 --------------------------RKQICELLLRKGANINEKTKEFLTPLHVASEKA 485
Db        520 VADFLIKAGADIELGCSTPLMEASQEGHLELVKYLLASGANVHATTATGDTALTYACENG 579

Qy        486 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR-------------------- 524
Db        580 HTDVADVLLQAGADLEHESEGGRTPLMKAARAGHLCTVQFLISKGANVNRATANNDHTVV 639
```

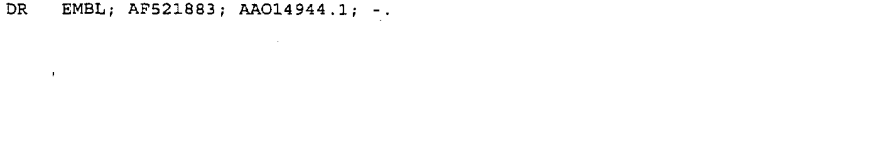


```
RESULT 45
Q8IWZ2
ID   Q8IWZ2      PRELIMINARY;      PRT;  2617 AA.
AC   Q8IWZ2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Multiple ankyrin repeats single KH domain protein isoform 2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Poulin F., Sonenberg N.;
RT   "Two overlapping reading frames in the second exon of the
RT   translational inhibitor 4E-BP3.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF521883; AAO14944.1; -.
```



```
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00023; ank; 25.
DR   Pfam; PF00013; KH; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00322; KH; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50084; KH_TYPE_1; 1.
SQ   SEQUENCE   2617 AA;  277173 MW;  2C93ED2B397C7CA6 CRC64;

  Query Match             8.6%;  Score 558.5;  DB 4;  Length 2617;
  Best Local Similarity   21.8%;  Pred. No. 7.4e-26;
  Matches  293;  Conservative  122;  Mismatches  418;  Indels  513;  Gaps   33;
```



```
Qy        594 ----------------------------RDIEGRQSTP---------------------- 603
                                           |::  | :|
Db        878 VFPEGEGDGSLPEDHFSELPQVDTILFKDNDVDDEQQSPPSAEQIDFVPVQPLSSPQCNF 937

Qy        604 ------------------------------------------------------------ 603

Db        938 SSDLGSNGTNSLELQKVSGNQQIVGQPQIAITGHDQGLLVQEPDGLMVATPAQTLTDTLD 997

Qy        604 ------------------------------------------------------------ 603

Db        998 DLIAAVSTRVPTGSNSSSQTTECLTPESCSQTTSNVASQSMPPVYPSVDIDAHTESNHDT 1057

Qy        604 -LHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLV-KHGAVVNV 661
               |  |       :|  |:   | :  :|| |  ||  | : ||  | |:|: | | :
Db       1058 ALTLACAGGHEELVSVLIARDAKIEHRDKKGFTPLILAATAGHVGVVEILLDKGGDIEAQ 1117

Qy        662 ADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD--TDIQDLLR- 718
              ::  | |||  | : |: |:  ||| ||:   :|    ||| |    |     |: ||
Db       1118 SERTKDTPLSLACSGGRQEVVDLLLARGANKEHRNVSDYTPLSLAASGGYVNIIKILLNA 1177

Qy        719 -----------GDAALLDAAKKGCLARVK-KLSSPDNVNCRDTQGRHSTPLHLAAGYNN 765
                        | : |: ||  | :  ||  |    ::| :    | :| | ||
Db       1178 GAEINSRTGSKLGISPLMLAAMNGHVPAVKLLLDMGSDINAQIETNR-NTALTLACFQGR 1236

Qy        766 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA--TDKWAFTPL 823
               ||   ||   |:|  : | || ||  ||| |: :|  :|:   | |||       | |
Db       1237 AEVVSLLLDRKANVEHRAKTGLTPLMEAASGGYAEVGRVLLDKGADVNAPPVPSSRDTAL 1296

Qy        824 HEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSAD---DVSALLTAAMPPSALPSC 880
              || ||  : | ||:  ||   ::|::| ||| | |     ||  || |
Db       1297 TIAADKGHYKFCELLIHRGAHIDVRNKKGNTPLWLASNGGHFDVVQLLVQA--------- 1347

Qy        881 YKPQVLNGVRSPGATADALSSGPSSP 906
                       ||  ||  :  :|
Db       1348 -----------GADVDAADNRKITP 1361
```

Search completed: March 25, 2004, 16:44:31
Job time : 86.812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:37:05 ; Search time 14.8376 Seconds
(without alignments)
4351.577 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6108	94.5	1166	1	TNK2_HUMAN	Q9h2k2 homo sapien
2	5103.5	79.0	1327	1	TNK1_HUMAN	O95271 homo sapien
3	892.5	13.8	4377	1	ANK3_HUMAN	Q12955 homo sapien
4	878.5	13.6	1880	1	ANK1_HUMAN	P16157 homo sapien
5	877	13.6	3924	1	ANK2_HUMAN	Q01484 homo sapien
6	874	13.5	1862	1	ANK1_MOUSE	Q02357 mus musculu
7	760.5	11.8	1059	1	Y379_HUMAN	O15084 homo sapien
8	571	8.8	768	1	YB23_HUMAN	Q9ulj7 homo sapien
9	485	7.5	1401	1	LATA_LATMA	P23631 latrodectus
10	453.5	7.0	1169	1	AFY1_MOUSE	Q810b6 mus musculu
11	448	6.9	1169	1	AFY1_HUMAN	Q9p2r3 homo sapien
12	431	6.7	603	1	V162_FOWPV	Q9j569 fowlpox vir
13	399.5	6.2	747	1	V222_FOWPV	Q9j513 fowlpox vir
14	390.5	6.0	832	1	ANR3_HUMAN	P57078 homo sapien
15	382.5	5.9	596	1	V024_FOWPV	Q9j5h7 fowlpox vir
16	380.5	5.9	668	1	V244_FOWPV	Q9j4z6 fowlpox vir
17	371	5.7	1432	1	DAK1_HUMAN	P53355 homo sapien
18	367.5	5.7	1150	1	ANS1_MOUSE	P59672 mus musculu
19	367	5.7	1442	1	DAK1_MOUSE	Q80ye7 mus musculu
20	357	5.5	587	1	ASB2_HUMAN	Q96q27 homo sapien
21	354.5	5.5	776	1	ANR5_HUMAN	Q9nu02 homo sapien
22	351	5.4	542	1	V155_FOWPV	Q9j5a7 fowlpox vir
23	351	5.4	775	1	ANR5_MOUSE	Q9d2j7 mus musculu
24	341.5	5.3	1134	1	ANS1_HUMAN	Q92625 homo sapien
25	330	5.1	1267	1	HMT1_HUMAN	Q9h9b1 homo sapien
26	327.5	5.1	525	1	V228_FOWPV	Q9j507 fowlpox vir
27	322	5.0	525	1	ASB3_MOUSE	Q9wv72 mus musculu
28	321	5.0	436	1	V245_FOWPV	Q9j4z5 fowlpox vir
29	318	4.9	592	1	V246_FOWPV	Q9j4z4 fowlpox vir
30	315.5	4.9	656	1	FEM1_CAEEL	P17221 caenorhabdi
31	312.5	4.8	568	1	PPGB_MOUSE	Q8vhq3 mus musculu
32	310	4.8	528	1	PPGA_HUMAN	Q96i34 homo sapien
33	309	4.8	524	1	PPGA_MOUSE	Q923m0 mus musculu
34	308	4.8	518	1	ASB3_HUMAN	Q9y575 homo sapien
35	307.5	4.8	568	1	PPGB_BOVIN	Q95n27 bos taurus
36	305	4.7	735	1	RN5A_MOUSE	Q05921 mus musculu
37	304	4.7	567	1	PPGB_HUMAN	Q96t49 homo sapien
38	303	4.7	429	1	AS10_HUMAN	Q8wxi3 homo sapien
39	300.5	4.6	692	1	ANR6_HUMAN	Q9y2g4 homo sapien
40	297	4.6	1210	1	BAT8_HUMAN	Q96kq7 homo sapien
41	295.5	4.6	360	1	Y045_METMA	Q8q0u0 methanosarc
42	291.5	4.5	741	1	RN5A_HUMAN	Q05823 homo sapien
43	291.5	4.5	950	1	ORP1_HUMAN	Q9bxw6 homo sapien
44	289	4.5	899	1	KBF2_HUMAN	Q00653 homo sapien
45	288	4.5	1263	1	BAT8_MOUSE	Q9z148 mus musculu

ALIGNMENTS

RESULT 1
TNK2_HUMAN
ID TNK2_HUMAN STANDARD; PRT; 1166 AA.
AC Q9H2K2; Q9H8F2; Q9HAS4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-
DE interacting ankyrin-related ADP-ribose polymerase 2) (Tankyrase-like
DE protein) (Tankyrase-related protein).
GN TNKS2 OR TNKL OR TANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=21072550; PubMed=11205898;
RA Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W.,
RA Glass B., Meese E.U.;
RT "Novel tankyrase-related gene detected with meningioma-specific
RT sera.";
RL Clin. Cancer Res. 7:113-119(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast carcinoma;
RX MEDLINE=21190090; PubMed=11294570;
RA Kuimov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,
RA Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;
RT "Cloning and characterization of TNKL, a member of tankyrase gene
RT family.";
RL Genes Immun. 2:52-55(2001).
RN [3]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC TISSUE=Liver;
RX MEDLINE=21264473; PubMed=11278563;
RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,
RA Eyre H.J., Sutherland G.R., Daly R.J.;
RT "Identification of a novel human tankyrase through its interaction
RT with the adaptor protein Grb14.";
RL J. Biol. Chem. 276:17172-17180(2001).
RN [4]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC TISSUE=Placenta;
RX MEDLINE=21443728; PubMed=11454873;
RA Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D.,
RA Morin G.B., Yaswen P., Campisi J.;
RT "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
RT rapid induction of cell death upon overexpression.";
RL J. Biol. Chem. 276:35891-35899(2001).
RN [5]
RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND
RP LNPEP/OTASE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21661461; PubMed=11802774;

```
RA   Sbodio J.I., Lodish H.F., Chi N.-W.;
RT   "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1
RT   (telomere-repeat-binding factor 1) and IRAP (insulin-responsive
RT   aminopeptidase).";
RL   Biochem. J. 361:451-459(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Yin Y., Gelmann E.P.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Tracey A.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 838-1151 FROM N.A.
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   FUNCTION, AND ADP-RIBOSYLATION.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC       subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC       nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS. Interacts with the
CC       cytoplasmic domain of LNPEP/Otase in SLC2A4/GLUT4-vesicles. Binds
CC       to the N-terminus of Grb14 and TRF1 with its ankyrin repeat
CC       region.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4-vesicles. Also found around the
CC       pericentriolar matrix of mitotic centromeres. During interphase, a
CC       small fraction of TNKS2 is found in the nucleus, associated with
CC       TRF1.
CC   -!- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,
CC       liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral
CC       blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,
CC       ovary, small intestine, colon, mammary gland, breast and breast
CC       carcinoma, and in common-type meningioma. Highly expressed in
CC       fetal liver, heart and brain.
CC   -!- PTM: ADP-ribosylated (-auto).
CC   -!- SIMILARITY: Belongs to the PARP family.
CC   -!- SIMILARITY: Contains 15 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF305081; AAG25674.1; ALT_INIT.
DR   EMBL; AF264912; AAG44694.1; -.
DR   EMBL; AF329696; AAK13463.1; -.
DR   EMBL; AF342982; AAK25811.1; -.
DR   EMBL; AF309033; AAK82330.1; -.
DR   EMBL; AF438201; AAL40795.1; -.
DR   EMBL; AL359707; CAC78760.1; -.
DR   EMBL; AK023746; BAB14665.1; ALT_INIT.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:15677; TNKS2.
DR   MIM; 607128; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 18.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.
FT   REPEAT      57    89       ANK 1.
FT   REPEAT      90   122       ANK 2.
FT   REPEAT     123   155       ANK 3.
FT   REPEAT     210   242       ANK 4.
FT   REPEAT     243   275       ANK 5.
FT   REPEAT     276   308       ANK 6.
FT   REPEAT     363   398       ANK 7.
FT   REPEAT     399   431       ANK 8.
FT   REPEAT     432   464       ANK 9.
FT   REPEAT     525   557       ANK 10.
FT   REPEAT     558   590       ANK 11.
FT   REPEAT     591   623       ANK 12.
FT   REPEAT     678   710       ANK 13.
FT   REPEAT     711   743       ANK 14.
FT   REPEAT     744   776       ANK 15.
FT   DOMAIN     873   936       SAM.
FT   DOMAIN    1023  1162       PARP.
FT   CONFLICT   331   337       KGHSLLQ -> QRPLVAA (IN REF. 1).
FT   CONFLICT   357   361       NFKHP -> IQAS (IN REF. 1).
FT   CONFLICT   966   966       Q -> P (IN REF. 8).
SQ   SEQUENCE   1166 AA;  126917 MW;  4C8B3B8D97CEF704 CRC64;

  Query Match             94.5%;  Score 6108;  DB 1;  Length 1166;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1166;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy        735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
                |||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166


RESULT 2
TNK1_HUMAN
ID   TNK1_HUMAN     STANDARD;      PRT;  1327 AA.
AC   O95271; O95272;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase I) (TNKS-1) (TRF1-
DE   interacting ankyrin-related ADP-ribose polymerase).
GN   TNKS OR TNKS1 OR TIN1 OR TINF1 OR PARPL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Testis;
RX   MEDLINE=99040105; PubMed=9822378;
RA   Smith S., Giriat I., Schmitt A., de Lange T.;
RT   "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
RL   Science 282:1484-1487(1998).
RN   [2]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=99454782; PubMed=10523501;
RA   Smith S., de Lange T.;
RT   "Cell cycle dependent localization of the telomeric PARP, tankyrase,
RT   to nuclear pore complexes and centrosomes.";
RL   J. Cell Sci. 112:3649-3656(1999).
RN   [3]
RP   FUNCTION, AND PHOSPHORYLATION.
RX   MEDLINE=20556282; PubMed=10988299;
RA   Chi N.-W., Lodish H.F.;
RT   "Tankyrase is a Golgi-associated mitogen-activated protein kinase
RT   substrate that interacts with IRAP in GLUT4 vesicles.";
RL   J. Biol. Chem. 275:38437-38444(2000).
RN   [4]
RP   FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC       subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC       nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
CC       the cytoplasmic domain of LNPEP/Otase in SLC2A4/GLUT4-vesicles.
CC       Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is
CC       also found at nuclear pore complexes and around the pericentriolar
CC       matrix of mitotic centromeres. During interphase, a small fraction
CC       of TNKS is found in the nucleus, associated with TRF1.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=O95271-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=O95271-2; Sequence=VSP_004538, VSP_004539;
CC         Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC   -!- PTM: Upon insulin-stimulation, phosphorylated on serine residues
CC       by MAPK kinases.
CC   -!- PTM: ADP-ribosylated (-auto).
CC   -!- SIMILARITY: Belongs to the PARP family.
CC   -!- SIMILARITY: Contains 15 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF082556; AAC79841.1; -.
DR   EMBL; AF082557; AAC79842.1; -.
DR   EMBL; AF082558; AAC79843.1; -.
DR   EMBL; AF082559; AAC79844.1; -.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:11941; TNKS.
DR   MIM; 603303; -.
DR   GO; GO:0000781; C:chromosome, telomeric region; IDA.
DR   GO; GO:0003950; F:NAD ADP-ribosyltransferase activity; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0007004; P:telomerase-dependent telomere maintenance; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 19.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 17.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
```

```
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
KW   Phosphorylation; Alternative splicing.
FT   REPEAT      215    247       ANK 1.
FT   REPEAT      248    280       ANK 2.
FT   REPEAT      281    313       ANK 3.
FT   REPEAT      368    400       ANK 4.
FT   REPEAT      401    433       ANK 5.
FT   REPEAT      434    466       ANK 6.
FT   REPEAT      521    556       ANK 7.
FT   REPEAT      557    589       ANK 8.
FT   REPEAT      590    622       ANK 9.
FT   REPEAT      683    715       ANK 10.
FT   REPEAT      716    748       ANK 11.
FT   REPEAT      749    781       ANK 12.
FT   REPEAT      836    868       ANK 13.
FT   REPEAT      869    901       ANK 14.
FT   REPEAT      902    934       ANK 15.
FT   DOMAIN     1030   1089       SAM.
FT   DOMAIN     1176   1327       PARP.
FT   DOMAIN        9     14       POLY-HIS.
FT   DOMAIN       27     34       POLY-PRO.
FT   DOMAIN      128    134       POLY-SER.
FT   DOMAIN      137    145       POLY-SER.
FT   VARSPLIC    641    643       EST -> GHS (in isoform 2).
FT                                /FTId=VSP_004538.
FT   VARSPLIC    644   1327       Missing (in isoform 2).
FT                                /FTId=VSP_004539.
FT   MUTAGEN    1184   1184       H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT                                WITH A-1291.
FT   MUTAGEN    1291   1291       E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT                                WITH A-1184.
SQ   SEQUENCE   1327 AA;  142010 MW;  E14DE985C710B957 CRC64;

  Query Match             79.0%;  Score 5103.5;  DB 1;  Length 1327;
  Best Local Similarity   79.2%;  Pred. No. 1.5e-292;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :  : ||      : |      | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||        ||  | ||| |||||||| ||||||    ||:|  |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              |||||||:||| |||| |||||||||||||||||||||:||  ||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:|||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:|| || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
```

```
              ||| ||||:|||||  |||||||||||| ||:|||||||| |||||  |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
              |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              ||||||||||||||||||||||||||||||||||:||| |||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||:|||||||||:|||||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
              :|:|||||||:||||||||||||||||:|||||||||||||||||||||||||||||:
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
              |||||||| ||||||||||||||||||||||||||||||||||||:|||||||||| :|
Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
              ||: |||  |||| |||:|:|||          : :| |  |:|| ||||||:|||:|
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
              :||:   :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
              :|||:|||||||||||||||||||: |||| |||||  :    ||||:||:|:|||:
Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
              |||||||||:|||||||:|||||||:::|||| |||| ||: ||:|||||||||| |
Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
              |||||||||:|||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
              |||||||||:|||||||||||||| |||||:|||||| |||||||| |||||||||
Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235
              |||||||||||||:||
Db       1299 QAYPEYLITYQIMKPE 1314
```

```
RESULT 3
ANK3_HUMAN
ID   ANK3_HUMAN     STANDARD;      PRT;  4377 AA.
AC   Q12955;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin 3 (ANK-3) (Ankyrin G).
GN   ANK3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain stem;
RX   MEDLINE=95138209; PubMed=7836469;
RA   Kordeli E., Lambert S., Bennett V.;
RT   "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RT   at the axonal initial segment and node of Ranvier.";
RL   J. Biol. Chem. 270:2352-2359(1995).
```

```
CC   -!- FUNCTION: Membrane-cytoskeleton linker. The neural-specific
CC       isoforms may participate in the maintenance/targeting of ion
CC       channels and cell adhesion molecules at the nodes of Ranvier and
CC       axonal initial segments.
CC   -!- SUBUNIT: Neural-specific isoforms may be a constituent of a
CC       neurofascin/NRCAM/ankyrin G complex.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=1;
CC         Comment=A number of isoforms are produced;
CC       Name=1; Synonyms=480-kDa isoform;
CC         IsoId=Q12955-1; Sequence=Displayed;
CC   -!- TISSUE SPECIFICITY: Expressed in brain and other tissues. Isoform
CC       1 is neural-specific.
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U13616; AAA64834.1; -.
DR   PIR; A55575; A55575.
DR   HSSP; P55273; 1BI8.
DR   Genew; HGNC:494; ANK3.
DR   MIM; 600465; -.
DR   GO; GO:0006605; P:protein targeting; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT   REPEAT       73    102       ANK 1.
FT   REPEAT      106    135       ANK 2.
FT   REPEAT      139    168       ANK 3.
FT   REPEAT      172    201       ANK 4.
FT   REPEAT      203    230       ANK 5.
FT   REPEAT      234    263       ANK 6.
FT   REPEAT      267    296       ANK 7.
FT   REPEAT      300    329       ANK 8.
FT   REPEAT      333    362       ANK 9.
FT   REPEAT      366    395       ANK 10.
FT   REPEAT      399    428       ANK 11.
FT   REPEAT      432    461       ANK 12.
FT   REPEAT      465    494       ANK 13.
FT   REPEAT      498    527       ANK 14.
FT   REPEAT      531    560       ANK 15.
FT   REPEAT      564    593       ANK 16.
FT   REPEAT      597    626       ANK 17.
FT   REPEAT      630    659       ANK 18.
FT   REPEAT      663    692       ANK 19.
FT   REPEAT      696    725       ANK 20.
FT   REPEAT      729    758       ANK 21.
FT   REPEAT      762    791       ANK 22.
FT   REPEAT      795    825       ANK 23.
FT   DOMAIN     1519   1898       SER-RICH.
FT   DOMAIN     4090   4174       DEATH.
SQ   SEQUENCE   4377 AA;  480399 MW;  F42379E55768B684 CRC64;

  Query Match             13.8%;  Score 892.5;  DB 1;  Length 4377;
  Best Local Similarity   30.7%;  Pred. No. 7.8e-44;
```

```
  Matches  277;  Conservative  126;  Mismatches  365;  Indels  135;  Gaps   21;

Qy       136 LHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD 195
Db        78 LHLASKEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQS 137

Qy       196 NWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEY 250
Db       138 QNGFTPLYMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDT 197

Qy       251 K---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQH 307
Db       198 KGKVRLPALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINVATLLLNR 257

Qy       308 GADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVC 367
Db       258 AAAVDFTARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVV 317

Qy       368 SLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL 427
Db       318 EMLLDRAA-PILSKTKN--------------------GLSPLHMATQGD------HLNC 349

Qy       428 EMVNFKH-------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
Db       350 VQLLLQHNVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTPLHI 406

Qy       481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
Db       407 ACKKNRIKVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRG 466

Qy       541 FTALQMG-----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVN 592
Db       467 ETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGASPN 526

Qy       593 CRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELL 652
Db       527 AATTSG--YTPLHLSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVANLL 584

Qy       653 VKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDT 711
Db       585 LQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQM 644

Qy       712 DIQDLL------------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLH 758
Db       645 DIATTLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNANVNLSNKSG--LTPLH 702

Qy       759 LAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
Db       703 LAAQEDRVNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKN 762

Qy       819 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLT-AA 871
Db       763 GYTPLHQAAQQGHTHIINVLLQNNASPNELTVNGNTALGIARRLGYISVVDTLKIVTEET 822

Qy       872 MPPSALPSCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSV 928
Db       823 MTTTTVTEKHKMNVPETMNEVLD------------MSDDEVRKANAPEMLSDG------- 863

Qy       929 VSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ-ITLDVLVEMGHKELK 987
Db       864 --------------------------------EYISDVEEGEDAMTGDTDKYLGPQDLK 890

Qy       988 EIG 990
Db       891 ELG 893


RESULT 4
ANK1_HUMAN
ID   ANK1_HUMAN     STANDARD;      PRT;  1880 AA.
AC   P16157;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
```

```
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN   ANK1 OR ANK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC   TISSUE=Hematopoietic;
RX   MEDLINE=90158830; PubMed=2137557;
RA   Lux S.E., John K.M., Bennett V.;
RT   "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT   structure with homology to tissue-differentiation and cell-cycle
RT   control proteins.";
RL   Nature 344:36-42(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90175370; PubMed=1689849;
RA   Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
RA   Cheung M.C., Kan Y.W., Palek J.;
RT   "cDNA sequence for human erythrocyte ankyrin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN   [3]
RP   VARIANT HS ILE-462.
RX   MEDLINE=96225450; PubMed=8640229;
RA   Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA   Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
RA   Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT   "Ankyrin-1 mutations are a major cause of dominant and recessive
RT   hereditary spherocytosis.";
RL   Nat. Genet. 13:214-218(1996).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements; bind to the erythrocyte membrane protein band 4.2, to
CC       Na-K ATPase, to the lymphocyte membrane protein GP85, and to the
CC       cytoskeletal proteins fodrin, tubulin, vimentin and desmin.
CC       Erythrocyte ankyrins also link spectrin (beta chain) to the
CC       cytoplasmic domain of the erythrocytes anion exchange protein;
CC       they retain most or all of these binding functions.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA
CC       MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1; Synonyms=2.1;
CC         IsoId=P16157-1; Sequence=Displayed;
CC       Name=2; Synonyms=2.2;
CC         IsoId=P16157-2; Sequence=VSP_000264, VSP_000265;
CC       Name=3;
CC         IsoId=P16157-3; Sequence=VSP_000266;
CC   -!- PTM: Regulated by phosphorylation.
CC   -!- PTM: Palmitoylated.
CC   -!- DISEASE: Defects in ANK1 are a cause of hereditary spherocytosis
CC       (HS) [MIM:182900]. Inheritance can be autosomal dominant or
CC       recessive.
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16609; CAA34610.1; -.
DR   EMBL; M28880; AAA51732.1; -.
DR   PIR; A35049; A35049.
DR   PIR; S08275; SJHUK.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:492; ANK1.
DR   MIM; 182900; -.
DR   GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;
KW   Polymorphism.
FT   INIT_MET      0      0
FT   DOMAIN        1    826       89 kDa DOMAIN
FT                                (ANION EXCHANGE PROTEIN BINDING DOMAIN).
FT   DOMAIN      827   1381       62 kDa DOMAIN (SPECTRIN BINDING DOMAIN).
FT   DOMAIN     1382   1880       55 kDa REGULATORY DOMAIN
FT                                (REGULATES THE BINDING OF ANKYRIN TO
FT                                SPECTRIN AND THE BAND 3 PROTEIN).
FT   REPEAT       43     72       ANK 1.
FT   REPEAT       76    105       ANK 2.
FT   REPEAT      109    138       ANK 3.
FT   REPEAT      142    171       ANK 4.
FT   REPEAT      173    200       ANK 5.
FT   REPEAT      204    233       ANK 6.
FT   REPEAT      237    266       ANK 7.
FT   REPEAT      270    299       ANK 8.
FT   REPEAT      303    332       ANK 9.
FT   REPEAT      336    365       ANK 10.
FT   REPEAT      369    398       ANK 11.
FT   REPEAT      402    431       ANK 12.
FT   REPEAT      435    464       ANK 13.
FT   REPEAT      468    497       ANK 14.
FT   REPEAT      501    530       ANK 15.
FT   REPEAT      534    563       ANK 16.
FT   REPEAT      567    596       ANK 17.
FT   REPEAT      600    629       ANK 18.
FT   REPEAT      633    662       ANK 19.
FT   REPEAT      666    695       ANK 20.
FT   REPEAT      699    728       ANK 21.
FT   REPEAT      732    761       ANK 22.
FT   REPEAT      765    794       ANK 23.
FT   DOMAIN     1402   1486       DEATH.
FT   VARSPLIC   1512   1873       Missing (in isoform 2).
FT                                /FTId=VSP_000264.
FT   VARSPLIC   1874   1874       H -> D (in isoform 2).
FT                                /FTId=VSP_000265.
FT   VARSPLIC   1849   1880       TVEGPLEDPSELEVDIDYFMKHSKDHTSTPNP -> ELRGS
FT                                GLQPDLIEGRKGAQIVKRASLKRGKQ (in isoform
FT                                3).
FT                                /FTId=VSP_000266.
FT   VARIANT      20     20       R -> T.
FT                                /FTId=VAR_000595.
FT   VARIANT     462    462       V -> I (in HS).
FT                                /FTId=VAR_000596.
FT   VARIANT     618    618       R -> H (in Brueggen).
FT                                /FTId=VAR_000597.
FT   VARIANT     749    749       V -> A.
FT                                /FTId=VAR_000598.
FT   VARIANT     844    844       D -> E.
FT                                /FTId=VAR_000599.
FT   VARIANT    1285   1285       E -> D.
FT                                /FTId=VAR_000601.
FT   VARIANT    1391   1391       S -> T.
FT                                /FTId=VAR_000600.
FT   VARIANT    1591   1591       D -> N (in Duesseldorf).
FT                                /FTId=VAR_000602.
FT   VARIANT    1698   1698       R -> D.
```

```
FT                              /FTId=VAR_000603.
FT   CONFLICT    229    229     A -> S (IN REF. 2).
FT   CONFLICT   1545   1545     V -> I (IN REF. 2).
SQ   SEQUENCE   1880 AA;  206145 MW;  1C5F5E7EFD1CD428 CRC64;

  Query Match             13.6%;  Score 878.5;  DB 1;  Length 1880;
  Best Local Similarity   25.8%;  Pred. No. 1.5e-43;
  Matches  292;  Conservative  157;  Mismatches  375;  Indels  307;  Gaps   28;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
Db         51 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 109

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
Db        110 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 169

Qy        225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
Db        170 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 210

Qy        276 VNCHASDGRKS-----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
Db        211 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETKTKD 270

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
Db        271 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYDAE- 329

Qy        378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437
Db        330 ----------ID--------------------------DIT---------------LD 336

Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
Db        337 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRVMELLLKTG 393

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQ 552
Db        394 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 453

Qy        553 QLLQE-----------------GISLGNSEADRQLLE-----------------AAKAG 577
Db        454 YLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAREG 513

Qy        578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
Db        514 HVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVRVAELLLERDAHPNAAGKNGLTP 571

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
Db        572 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQVEVARSLLQYGGSANAE 630

Qy        696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
Db        631 SVQGVTPLHLAAQEGHAEMVALL-------------------LSKQANGNLGNKSG--L 668

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
Db        669 TPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 728

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL--------------VS 860
Db        729 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDGTTPLAIAKRLGYISVTDVLKVV 788

Qy        861 ADDVSALLTAAMPPSALPSC-----------------------------------YKPQ 884
Db        789 TDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISFKAERRDSRDVDEEKELLDFVPK 848

Qy        885 VLNGVRSPGATADALSSGPSSPSSLS-----AASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        849 LDQVVESPAI--------PRIPCAMPETVVIRSEEQEQASKEYDEDSLIPSSPATETSDN 900

Qy        940 LEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHK 999
Db        901 ISPVASPVHTGFLVSFM----------------------------VDARGGSMRGSRHN 931

Qy       1000 LIKGVE------------RLISGQQGLNPYLTLNTSGSGTILIDLSP------------ 1034
Db        932 GLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEGLASRIIALGPTGAQFLSPVIVE 991

Qy       1035 ---------DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 1076
Db        992 IPHFASHGRGDRELVVLRSENGSVWKEHR---------SRYGESYLDQILN 1033


RESULT 5
ANK2_HUMAN
ID   ANK2_HUMAN     STANDARD;      PRT;  3924 AA.
AC   Q01484; Q01485;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN   ANK2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC   TISSUE=Brain stem;
RX   MEDLINE=91302466; PubMed=1830053;
RA   Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT   "Isolation and characterization of cDNAs encoding human brain
RT   ankyrins reveal a family of alternatively spliced genes.";
RL   J. Cell Biol. 114:241-253(1991).
RN   [2]
RP   REVISIONS.
RA   Carpenter S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain stem;
RX   MEDLINE=94075409; PubMed=8253844;
RA   Chan W., Kordeli E., Bennett V.;
RT   "440-kD ankyrinB: structure of the major developmentally regulated
RT   domain and selective localization in unmyelinated axons.";
RL   J. Cell Biol. 123:1463-1473(1993).
RN   [4]
RP   SEQUENCE OF 463-495 FROM N.A.
RX   MEDLINE=92009921; PubMed=1833308;
RA   Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA   Lux S.E., Ward D.C., Forget B.G.;
RT   "Isolation and chromosomal localization of a novel nonerythroid
RT   ankyrin gene.";
RL   Genomics 10:858-866(1991).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements. Also bind to cytoskeletal proteins.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=Q01484-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q01484-2; Sequence=VSP_000267, VSP_000268;
CC       Name=3;
CC         IsoId=Q01484-3; Sequence=VSP_000268;
CC   -!- TISSUE SPECIFICITY: Plasma membrane of neurons as well as glial
CC       cells throughout the brain.
CC   -!- PTM: Phosphorylated at multiple sites by different protein kinases
CC       and each phosphorylation event regulates the protein's structure
CC       and function (Potential).
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X56957; CAA40278.1; -.
DR   EMBL; X56958; CAA40279.2; -.
DR   EMBL; Z26634; CAB42644.1; -.
DR   EMBL; M37123; AAA62828.1; -.
DR   PIR; S37431; S37431.
DR   HSSP; P42771; 1DC2.
DR   Genew; HGNC:493; ANK2.
DR   MIM; 106410; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Phosphorylation.
FT   REPEAT       63     92       ANK 1.
FT   REPEAT       96    125       ANK 2.
FT   REPEAT      129    158       ANK 3.
FT   REPEAT      162    191       ANK 4.
FT   REPEAT      193    220       ANK 5.
FT   REPEAT      232    261       ANK 6.
FT   REPEAT      265    294       ANK 7.
FT   REPEAT      298    327       ANK 8.
FT   REPEAT      331    360       ANK 9.
FT   REPEAT      364    393       ANK 10.
FT   REPEAT      397    426       ANK 11.
FT   REPEAT      430    459       ANK 12.
FT   REPEAT      463    492       ANK 13.
FT   REPEAT      496    525       ANK 14.
FT   REPEAT      529    558       ANK 15.
FT   REPEAT      562    591       ANK 16.
FT   REPEAT      595    624       ANK 17.
FT   REPEAT      628    657       ANK 18.
FT   REPEAT      661    690       ANK 19.
FT   REPEAT      694    723       ANK 20.
FT   REPEAT      727    756       ANK 21.
FT   REPEAT      760    789       ANK 22.
FT   REPEAT      793    822       ANK 23.
FT   DOMAIN     1773   1950       REPEAT-RICH REGION.
FT   REPEAT     1773   1784       REPEAT A.
FT   REPEAT     1785   1796       REPEAT A.
FT   REPEAT     1797   1808       REPEAT A.
FT   REPEAT     1809   1820       REPEAT A.
FT   REPEAT     1821   1832       REPEAT A.
FT   REPEAT     1833   1844       REPEAT A.
FT   REPEAT     1845   1856       REPEAT A.
FT   REPEAT     1857   1867       REPEAT A (APPROXIMATE).
FT   REPEAT     1868   1879       REPEAT A.
FT   REPEAT     1880   1891       REPEAT A.
FT   REPEAT     1892   1902       REPEAT A (APPROXIMATE).
FT   REPEAT     1903   1914       REPEAT A.
FT   REPEAT     1915   1926       REPEAT A.
FT   REPEAT     1927   1938       REPEAT A.
FT   REPEAT     1939   1950       REPEAT A.
FT   DOMAIN     3536   3620       DEATH.
FT   VARSPLIC   1039   1039       Q -> QFLGKLHLPTAPPPLNEGESLVSRILQLGPPGTK
FT                                (in isoform 2).
FT                                /FTId=VSP_000267.
FT   VARSPLIC   1444   3528       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_000268.
FT   CONFLICT    475    476       GQ -> PE (IN REF. 4).
FT   CONFLICT    971    971       I -> S (IN REF. 1).
FT   CONFLICT   3581   3582       QY -> HA (IN REF. 1).
FT   CONFLICT   3586   3586       I -> Y (IN REF. 1).
SQ   SEQUENCE   3924 AA;  430337 MW;  52AC496C428E29D2 CRC64;

  Query Match             13.6%;  Score 877;  DB 1;  Length 3924;
  Best Local Similarity   29.7%;  Pred. No. 5.4e-43;
  Matches  293;  Conservative  137;  Mismatches  402;  Indels  156;  Gaps   23;

Qy       105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
Db        38 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 95

Qy       164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
Db        96 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 155

Qy       224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
Db       156 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 213

Qy       276 VNCHASDGRKS-----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
Db       214 -NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 272

Qy       325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
Db       273 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 331

Qy       385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
Db       332 --------------------GLSPLHMAAQGDHVECVKHL----LQHKAPVDDVTLDYL 366

Qy       440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
Db       367 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAS 423

Qy       500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN-----ENVQQL 554
Db       424 IQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCL 483

Qy       555 LQEGI-----------------SLGNSEADRQLLE-----------------AAKAGDV 579
Db       484 LRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTNGYTPLHISAREGQV 543

Qy       580 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
Db       544 DVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHV 602

Qy       640 ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDG 699
Db       603 AAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNIVTKQG 662

Qy       700 NTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLH 758
Db       663 VTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----------------------TSLH 700

Qy       759 LAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
Db       701 LAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYGNVKMVNFLLKQGANVNAKTKN 760

Qy       819 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLTAAM 872
Db       761 GYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALAIAKRLGYISVVDTLKVVTEEV 820

Qy       873 PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSEL------- 925
Db       821 TTTTTTITEK----HKLNVPETMTEVL--------DVSDEEGDDTMTGDGGEYLRPEDLK 868

Qy       926 ----SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RNLGLEHLMDIFEREQITLDVLV 979
```

```
Db        869 ELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRSHTLSHAS--YLRDSAVMDDSV 926

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERL 1007
Db        927 VIPSHQVSTLAKEAERNSYRLSWGTENL 954


RESULT 6
ANK1_MOUSE
ID   ANK1_MOUSE     STANDARD;      PRT;  1862 AA.
AC   Q02357;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ankyrin 1 (Erythrocyte ankyrin).
GN   ANK1 OR ANK-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=92345717; PubMed=1386265;
RA   White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT   "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT   regulatory domain.";
RL   Mamm. Genome 3:281-285(1992).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements; bind to the erythrocyte membrane protein band 4.2, to
CC       Na-K ATPase, to the lymphocyte membrane protein GP85, and to the
CC       cytoskeletal proteins fodrin, tubulin, vimentin and desmin.
CC       Erythrocyte ankyrins also link spectrin (beta chain) to the
CC       cytoplasmic domain of the erythrocytes anion exchange protein;
CC       they retain most or all of these binding functions.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC       PLASMA MEMBRANE.
CC   -!- PTM: Regulated by phosphorylation (By similarity).
CC   -!- PTM: Acylated by palmitic acid group(s) (By similarity).
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M84756; AAA37236.1; -.
DR   PIR; I49502; I49502.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:88024; Ank1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
FT   DOMAIN        1    827       89 kDa DOMAIN (ANION EXCHANGE PROTEIN
FT                                BINDING DOMAIN).
FT   DOMAIN      828   1386       62 kDa DOMAIN (SPECTRIN BINDING
FT                                DOMAIN).
FT   DOMAIN     1387   1862       55 kDa REGULATORY DOMAIN (REGULATES
FT                                THE BINDING OF ANKYRIN TO SPECTRIN
FT                                AND THE BAND 3 PROTEIN).
FT   REPEAT       40     69       ANK 1.
FT   REPEAT       73    102       ANK 2.
FT   REPEAT      106    135       ANK 3.
FT   REPEAT      139    168       ANK 4.
FT   REPEAT      170    197       ANK 5.
FT   REPEAT      201    230       ANK 6.
FT   REPEAT      234    263       ANK 7.
FT   REPEAT      267    296       ANK 8.
FT   REPEAT      300    329       ANK 9.
FT   REPEAT      333    362       ANK 10.
FT   REPEAT      366    395       ANK 11.
FT   REPEAT      399    428       ANK 12.
FT   REPEAT      432    461       ANK 13.
FT   REPEAT      465    494       ANK 14.
FT   REPEAT      498    527       ANK 15.
FT   REPEAT      531    560       ANK 16.
FT   REPEAT      564    593       ANK 17.
FT   REPEAT      597    626       ANK 18.
FT   REPEAT      630    659       ANK 19.
FT   REPEAT      663    692       ANK 20.
FT   REPEAT      696    725       ANK 21.
FT   REPEAT      729    758       ANK 22.
FT   REPEAT      762    791       ANK 23.
FT   DOMAIN     1399   1483       DEATH.
SQ   SEQUENCE   1862 AA;  204242 MW;  AE6B85B5B29001E5 CRC64;

  Query Match           13.5%;  Score 874;  DB 1;  Length 1862;
  Best Local Similarity  29.7%;  Pred. No. 2.7e-43;
  Matches  250;  Conservative  132;  Mismatches  281;  Indels  178;  Gaps   21;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
Db         48 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 106

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
Db        107 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 166

Qy        225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
Db        167 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 207

Qy        276 VNCHASDGRKS------------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
Db        208 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETRTKD 267

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
Db        268 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYNAE- 326

Qy        378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437
Db        327 ----------ID--------------------------DIT----------------LD 333

Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
Db        334 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHIRVMELLLKTG 390

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM----GNENVQQ 553
Db        391 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 450

Qy        554 LLQEGISLGNSEAD----------------RQLLE----------------AAKAG 577
Db        451 YLLQNKAKANAKADDQTPLHCAARIGHTGMVKLLLENGASPNLATTAGHTPLHTAAREG 510

Qy        578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
Db        511 HVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVRLAELLLEHDAHPNAAGKNGLTP 568
```

```
Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
              || |  : : :: :||:  |   : :  |  :|||| || : : |: : |||:|     :
Db        569 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQIEVARSLLQYGGSANAE 627

Qy        696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
              :  | ||| |  ::| |::  ||                    ||   | |  :  |
Db        628 SVQGVTPLHLAAQEGHTEMVALL-------------------LSKQANGNLGNKSG--L 665

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
              ||||| :   :: ||: |::||  |:|  : |  ||| |: ||:: :   |::: | |||
Db        666 TPLHLVSQEGHVLVADVLIKHGVTVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 725

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALL 868
                |  ::|||:|||:| | :  ||| :|| |   :  | ||| :      :|  ||  ::
Db        726 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNGTTPLAIAKRLGYISVTDVLKVV 785

Qy        869 T 869
              |
Db        786 T 786
```

```
RESULT 7
Y379_HUMAN
ID   Y379_HUMAN     STANDARD;      PRT;  1059 AA.
AC   O15084;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical ANK-repeat protein KIAA0379 (Fragment).
GN   KIAA0379.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97349984; PubMed=9205841;
RA   Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VII.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 4:141-150(1997).
CC   -!- SIMILARITY: Contains at least 27 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB002377; BAA20833.2; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 28.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 28.
DR   PROSITE; PS50088; ANK_REPEAT; 24.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT       46     75       ANK 1.
FT   REPEAT       79    108       ANK 2.
FT   REPEAT      112    141       ANK 3.
FT   REPEAT      145    174       ANK 4.
FT   REPEAT      178    207       ANK 5.
FT   REPEAT      211    240       ANK 6.
FT   REPEAT      244    273       ANK 7.
FT   REPEAT      277    307       ANK 8.
FT   REPEAT      311    340       ANK 9.
FT   REPEAT      344    373       ANK 10.
FT   REPEAT      377    406       ANK 11.
FT   REPEAT      410    439       ANK 12.
FT   REPEAT      443    472       ANK 13.
FT   REPEAT      476    506       ANK 14.
FT   REPEAT      510    540       ANK 15.
FT   REPEAT      555    584       ANK 16.
FT   REPEAT      588    617       ANK 17.
FT   REPEAT      622    651       ANK 18.
FT   REPEAT      658    687       ANK 19.
FT   REPEAT      691    720       ANK 20.
FT   REPEAT      724    753       ANK 21.
FT   REPEAT      761    790       ANK 22.
FT   REPEAT      793    823       ANK 23.
FT   REPEAT      828    857       ANK 24.
FT   REPEAT      861    891       ANK 25.
FT   REPEAT      895    924       ANK 26.
FT   REPEAT      931    960       ANK 27.
SQ   SEQUENCE   1059 AA;  113465 MW;  C1F55E6CFE494770 CRC64;
```

```
  Query Match           11.8%;  Score 760.5;  DB 1;  Length 1059;
  Best Local Similarity  27.6%;  Pred. No. 5.7e-37;
  Matches  275;  Conservative  129;  Mismatches  357;  Indels  235;  Gaps   33;

Qy        102 LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
              | :|  ||| : |: |:   | || :|    | |||| ||  |  :::| |: :|| | |
Db         18 LVQAIFNGDPDEVRALIFKKEDVNFQD--NEKRTPLHAAAYLGDAEIIELLILSGARVNA 75

Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGKIDVCIVLLQ 219
              :|   | ||| | :   | | :||:| || |||| ||  |||| ||
Db         76 KDSKWLTPLHRAVASCSEEAVQVLLKHSADVNARDKNWQ-TPLHIAA------------ 121

Qy        220 HGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCH 279
                                  : |||                 |    | ||  | :
Db        122 ---------------------ANKAV-------------------KCAEALVPLLSNVN 140

Qy        280 ASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVK 339
              ||     | || ||      ::|:|||  ||:::| || |   :| |   || || :|||
Db        141 VSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFDKKDRRAIHWAAYMGHIEVVKLLVS 200

Qy        340 HGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKER 399
              ||| |    |   :|||| |||   : |   ||  | |     | :  : : :|
Db        201 HGAEVTCKDKKSYTPLHAAASSGMISVVKYLLDLGVDMNEPNAYGNTPLHVAC------- 253

Qy        400 LAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETALHCAAASPYPKRKQIC- 457
                : |  ::          : : :   :|| |: :   | || |||| :     :|
Db        254 ----YNGQDVV----------VNELIDCGAIVNQKNEKGF-TPLHFAAASTH---GALCL 295

Qy        458 ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 517
              |||:  ||::| |:|:  ||||: :        : :::  | ::  |  | | || ||
Db        296 ELLVGNGADVNMKSKDGKTPLHMTALHGRFSRSQTIIQSGAVIDCEDKNGNTPLHIAARY 355

Qy        518 GHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQQLLQEGISLGN-SEADRQLL 571
              ||       |:: | |      : |   | :      ::  ::||  |  :    :  | |
Db        356 GHELLINTLITSGADTAKRGIHGMFPLHLAALSGFSDCCRKLLSSGFDIDTPDDFGRTCL 415

Qy        572 EAAKA-GDVETVKKL-----------------------CTVQ----------SVNCRDI 596
              || | |::| :  |                       |  |          |||  |:
Db        416 HAAAAGGNLECLNLLLNTGADFNKKDKFGRSPLHYAAANCNYQCLFALVGSGASVN--DL 473

Qy        597 EGRQSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 655
              :  |  ||||:||  :     :||||:: |:   :|| |   :| | : :|||    :|:
Db        474 DERGCTPLHYAATSDTDGKCLEYLLRNDANPGIRDKQGYNAVHYSAAYGHRLCLQLIASE 533

Qy        656 GAV-----VNVADLWK-------FTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
                :       :  |:              :||| ||  | ::  ::|:|   |   :| | |||
Db        534 TPLDVLMETSGTDMLSDSDNRATISPLHLAAYHGHHQALEVLVQSLLDLDVRNSSGRTPL 593

Qy        704 DLVK-DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAG 762
              ||    |  :  |:|    | :   |  | |                    ||:| ||
Db        594 DLAAFKGHVECVDVLINQGASI-LVKDYILKR-------------------TPIHAAAT 632

Qy        763 YNNLEVAEYLLQHGADVNA---QDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 819
```

```
Db        633 NGHSECLRLLIGNAEPQNAVDIQDGNGQTPLMLSVLNGHTDCVYSLLNKGANVDAKDKWG 692

Qy        820 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA----DDVSALLTAAMPPS 875

Db        693 RTALHRGAVTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAACGHIGVLGALLQSAASMD 752

Qy        876 ALP-------------SCYKPQVLNG--------------VRSPGATADALSSGPSSPSS 908

Db        753 ANPATADNHGYTALHWACY-----NGHETCVELLLEQEVFQKTEGNAFSPLHCAVINDNE 807

Qy        909 LSAASSLDNLSGSFSELSSVVSSSGTEGASSL----------------EKKEVPGVDFS 951

Db        808 GAAEMLIDTLG------ASIVNATDSKGRTPLHAAAFTDHVECLQLLLSHNAQVNSVD-- 859

Qy        952 ITQFVRNLGLEHLMDIFEREQI-TLDVLVEMGHKEL 986

Db        860 ------STGKTPLMMAAENGQTNTVEMLVSSASAEL 889


RESULT 8
YB23_HUMAN
ID   YB23_HUMAN     STANDARD;      PRT;   768 AA.
AC   Q9ULJ7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein KIAA1223 (Fragment).
GN   KIAA1223.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K.-I., Kikuno R., Hirosawa M., Nomura N.,
RA   Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
CC   -!- SIMILARITY: Contains at least 14 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB033049; BAA86537.1; -.
DR   HSSP; P42771; 1DC2.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 14.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT        1     11       ANK 1.
FT   REPEAT       15     44       ANK 2.
FT   REPEAT       48     82       ANK 3.
FT   REPEAT       86    115       ANK 4.
FT   REPEAT      119    148       ANK 5.
FT   REPEAT      152    181       ANK 6.
FT   REPEAT      185    214       ANK 7.
FT   REPEAT      218    247       ANK 8.
FT   REPEAT      251    280       ANK 9.
FT   REPEAT      284    313       ANK 10.
FT   REPEAT      317    346       ANK 11.
FT   REPEAT      350    379       ANK 12.
FT   REPEAT      383    412       ANK 13.
FT   REPEAT      416    446       ANK 14.
SQ   SEQUENCE   768 AA;  82819 MW;  2913B69BE2DFE06D CRC64;

  Query Match             8.8%;  Score 571;  DB 1;  Length 768;
  Best Local Similarity   26.1%;  Pred. No. 5e-26;
  Matches  207;  Conservative  104;  Mismatches  255;  Indels  228;  Gaps   20;

Qy        184 LLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAK 243

Db          2 LLQHGAEVNKADNEGRTALIAAAYMGHREIVEHLLDHGAEVNHEDVDGRTALSVA----- 56

Qy        244 AVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQL 303

Db         57 ------------------------ALCVP------ASKGHAS-------------VVSL 72

Qy        304 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 363

Db         73 LIDRGAEVDHCDKDGMTPLLVAAYEGHVDVVDLLLEGGADVDHTDNNGRTPLLAAASMGH 132

Qy        364 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ-AAREADVTRIK 422

Db        133 ASVVNTLLFWGA----------AVD----------SIDSEGRTVLSIASAQGNVEVVR 170

Qy        423 KHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVAS 482

Db        171 TLLDRGLDENHRDDAGWTPLHMAA---FEGHRLICEALIEQGARTNEIDNDGRIPFILAS 227

Qy        483 EKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFT 542

Db        228 QEGHYDCVQILLENKSNIDQRGYDGRNALRVAALEGHRDIVELLFSHGAD---------- 277

Qy        543 ALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQST 602

Db        278 ----------------------------------------------VNCKDADGRPT- 288

Qy        603 PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 662

Db        289 -LYILALENQLTMAEYFLENGANVEASDAEGRTALHVSCWQGHMEMVQVLIAYHADVNAA 347

Qy        663 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAA 722

Db        348 DNEKRSALQSAAWQGHVKVVQLLIEHG------------------------------A 375

Qy        723 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 782

Db        376 VVD----------------------HTCNQGATALCIAAQEGHIDVVQVLLEHGADPNHA 413

Qy        783 DKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHG 842

Db        414 DQFGRTAMRVAAKNGHSQIIKLLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKV---- 467

Qy        843 ADPTLK-NQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSS 901

Db        468 QSLTIKSNSSGST-----GGGD--------MQPS----------LRGL--PNGPTHAFSS 502

Qy        902 GPSSPSSL--SAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNL 959

Db        503 PSESPDSTVDRQKSSLSNNSLKSSKNSSLRTTSSTATAQT-----VPIDSFHNLSFTEQI 557

Qy        960 GLEHLMDIFEREQI 973

Db        558 QQHSLPRSRSRQSI 571


RESULT 9
LATA_LATMA
ID   LATA_LATMA     STANDARD;      PRT;  1401 AA.
AC   P23631; Q25328;
DT   01-NOV-1991 (Rel. 20, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-latrotoxin precursor.
OS   Latrodectus mactans (Black widow spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Araneoidea; Theridiidae; Latrodectus.
OX   NCBI_TaxID=6924;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tredecimguttatus; TISSUE=Venom gland;
RX   MEDLINE=91031994; PubMed=1977615;
RA   Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;
RT   "Cloning and structure of cDNA encoding alpha-latrotoxin from black
RT   widow spider venom.";
RL   FEBS Lett. 270:127-131(1990).
RN   [2]
RP   PARTIAL SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=91362695; PubMed=1888339;
RA   Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;
RT   "Structure of tryptic fragments of a neurotoxin from black widow
RT   spider venom.";
RL   Bioorg. Khim. 17:437-441(1991).
CC   -!- FUNCTION: Binds to high-affinity receptors that are localized in
CC       the presynaptic plasma membrane of the nerve terminal. Binding of
CC       toxin results in calcium influx into nerve terminals, activation
CC       of synaptic vesicle exocytosis, and ATP depletion.
CC   -!- PTM: Alpha-latrotoxin may be processed at its C-terminus.
CC   -!- SIMILARITY: Contains 21 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X55009; CAA38753.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 20.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 20.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Toxin; Neurotoxin; Presynaptic neurotoxin; Repeat; ANK repeat; Signal.
FT   SIGNAL        1     20
FT   CHAIN        21   1401       ALPHA-LATROTOXIN.
FT   REPEAT      490    521       ANK 1.
FT   REPEAT      525    554       ANK 2.
FT   REPEAT      559    589       ANK 3.
FT   REPEAT      593    622       ANK 4.
FT   REPEAT      626    656       ANK 5.
FT   REPEAT      660    690       ANK 6.
FT   REPEAT      695    723       ANK 7.
FT   REPEAT      729    758       ANK 8.
FT   REPEAT      762    791       ANK 9.
FT   REPEAT      795    824       ANK 10.
FT   REPEAT      828    857       ANK 11.
FT   REPEAT      862    891       ANK 12.
FT   REPEAT      895    924       ANK 13.
FT   REPEAT      928    957       ANK 14.
FT   REPEAT      971   1003       ANK 15.
FT   REPEAT     1004   1033       ANK 16.
FT   REPEAT     1035   1064       ANK 17.
FT   REPEAT     1068   1097       ANK 18.
FT   REPEAT     1101   1131       ANK 19.
FT   REPEAT     1137   1166       ANK 20.
FT   REPEAT     1170   1199       ANK 21.
SQ   SEQUENCE   1401 AA;  156856 MW;  B38A22083C142A98 CRC64;

  Query Match             7.5%;  Score 485;  DB 1;  Length 1401;
  Best Local Similarity   23.5%;  Pred. No. 1.4e-20;

  Matches  191;  Conservative  100;  Mismatches  302;  Indels  220;  Gaps   22;

Qy        100 RELFEACRNGD-----VERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYL--L 152
Db        456 RDLYNAASNPDSAVGFKEFTK--LNYDGANIRATFDHGRTVFHAAAKSGNDKIMFGLTFL 513

Qy        153 QNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGKI 211
Db        514 AKSTELNQPDKKGYTPIHVAADSGNAGIVNLLIQRGVSINSKTYHFLQTPLHLAAQRGFV 573

Qy        212 DVCIVLLQHGAEPTI----RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKM 267
Db        574 TTFQRLME---SPEININERDKDGFTP--------------------LHYAIRGGERIL 609

Qy        268 MALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAKDKGDLVPLHNAC 326
Db        610 EAFLNQISIDVNAKSNTGLTPFHLAIIKNDWPVASTLLGSKKVDINAVDENNITALHYAA 669

Qy        327 SYGHYEVTE------------------------LLVKH-----------GACVNAMDLW 350
Db        670 ILGYLETTKQLINLKEINANVVSSPGLLSALHYAILYKHDDVASFLMRSSNVNVNLKALG 729

Qy        351 QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL 410
Db        730 GITPLHLAVIQGRKQILSLMFDIG---------------------------------- 753

Qy        411 QAAREADVTRIKKHLSLEMVNFKHPQTHE--TALHCAAASPYPKRKQICELLLRKGANIN 468
Db        754 ------------------VNIEQ-KTDEKYTPLHLAAMSKYP---ELIQILLDQGSNFE 790

Qy        469 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 528
Db        791 AKTNSGATPLHLATFKGKSQAALILLNNEVNWRDTDENGQMPIHGAAMTGLLDVAQAIIS 850

Qy        529 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 588
Db        851 --IDATVV---------------------------------------------------- 856

Qy        589 QSVNCRDIEGRQS-TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 647
Db        857 ------DIEDKNSDTPLNLAAQNSHIDVIKYFIDQGADINTRNKKGLAPLLAFSKKGNLD 910

Qy        648 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
Db        911 MVKYLFDKNANVYIADNDGMNFFYYAVQNGHLNIVKYAMSEKDKFEWSNTDNN------R 964

Qy        708 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHST--PLHLAAGYNN 765
Db        965 RDECPNEECAISHFAVCDAVQFDRIEIVKYFVG--------TLGNFAICGPLHQAARYGH 1016

Qy        766 LEVAEYLLQH---GADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTP 822
Db       1017 LDIVKYLVEEEFLSVDGSKTD----TPLCYASENGHFTVVQYLVSNGAKVNHDCGNGMTA 1072

Qy        823 LHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 855
Db       1073 IDKAITKNHLQVVQFLAANGVDFRRKNSRGTTP 1105


RESULT 10
AFY1_MOUSE
ID   AFY1_MOUSE     STANDARD;      PRT;  1169 AA.
AC   Q810B6; O54807; Q80TG6; Q80UH8;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin repeats- and FYVE domain-containing protein 1 (Ankyrin repeats
DE   hooked to a zinc finger motif).
GN   ANKFY1 OR ANKHZN OR KIAA1255.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```

```
RN   [1]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RX   MEDLINE=99194578; PubMed=10092534;
RA   Ito K., Ishii N., Miyashita A., Tominaga K., Kuriyama H., Maruyama H.,
RA   Shirai M., Naito M., Arakawa M., Kuwano R.;
RT   "Molecular cloning of a novel 130-kDa cytoplasmic protein, Ankhzn,
RT   containing ankyrin repeats hooked to a zinc finger motif.";
RL   Biochem. Biophys. Res. Commun. 257:206-213(1999).
RN   [2]
RP   REVISIONS, AND SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA   Maruyama H., Kuriyama H., Ishii N., Ito K., Odani S., Kuwano R.;
RT   "Genomic organization, alternative splicing, and promoter assay of the
RT   mouse Ankhzn gene.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 994-1169 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22579291; PubMed=12693553;
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA   Nakajima D., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT   II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT   cDNAs identified by screening of terminal sequences of cDNA clones
RT   randomly sampled from size-fractionated libraries.";
RL   DNA Res. 10:35-48(2003).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Also associated with endosomes.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q810B6-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q810B6-2; Sequence=VSP_007917, VSP_007918;
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC   -!- SIMILARITY: Contains 21 ANK repeats.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB011370; BAA24980.2; -.
DR   EMBL; AB098329; BAC67211.1; -.
DR   EMBL; AB098157; BAC67389.1; -.
DR   EMBL; AB098135; BAC67389.1; JOINED.
DR   EMBL; AB098136; BAC67389.1; JOINED.
DR   EMBL; AB098137; BAC67389.1; JOINED.
DR   EMBL; AB098138; BAC67389.1; JOINED.
DR   EMBL; AB098139; BAC67389.1; JOINED.
DR   EMBL; AB098140; BAC67389.1; JOINED.
DR   EMBL; AB098141; BAC67389.1; JOINED.
DR   EMBL; AB098142; BAC67389.1; JOINED.
DR   EMBL; AB098143; BAC67389.1; JOINED.
DR   EMBL; AB098144; BAC67389.1; JOINED.
DR   EMBL; AB098145; BAC67389.1; JOINED.
DR   EMBL; AB098146; BAC67389.1; JOINED.
DR   EMBL; AB098147; BAC67389.1; JOINED.
DR   EMBL; AB098148; BAC67389.1; JOINED.
DR   EMBL; AB098149; BAC67389.1; JOINED.
DR   EMBL; AB098150; BAC67389.1; JOINED.
DR   EMBL; AB098151; BAC67389.1; JOINED.
DR   EMBL; AB098152; BAC67389.1; JOINED.
DR   EMBL; AB098153; BAC67389.1; JOINED.
DR   EMBL; AB098154; BAC67389.1; JOINED.
DR   EMBL; AB098155; BAC67389.1; JOINED.
DR   EMBL; AB098156; BAC67389.1; JOINED.
DR   EMBL; AB098143; BAC67388.1; -.
DR   EMBL; AB098135; BAC67388.1; JOINED.
DR   EMBL; AB098136; BAC67388.1; JOINED.
DR   EMBL; AB098137; BAC67388.1; JOINED.
DR   EMBL; AB098138; BAC67388.1; JOINED.
DR   EMBL; AB098139; BAC67388.1; JOINED.
DR   EMBL; AB098140; BAC67388.1; JOINED.
DR   EMBL; AB098141; BAC67388.1; JOINED.
DR   EMBL; AB098142; BAC67388.1; JOINED.
DR   EMBL; AK122479; BAC65761.1; -.
DR   MGD; MGI:1337008; Ankfy1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF00023; ank; 21.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF01363; FYVE; 1.
DR   SMART; SM00248; ANK; 20.
DR   SMART; SM00225; BTB; 1.
DR   SMART; SM00064; FYVE; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
KW   ANK repeat; Repeat; Zinc-finger; Alternative splicing.
FT   DOMAIN       68    130       BTB.
FT   REPEAT      217    247       ANK 1.
FT   REPEAT      255    284       ANK 2.
FT   REPEAT      288    317       ANK 3.
FT   REPEAT      322    362       ANK 4.
FT   REPEAT      366    396       ANK 5.
FT   REPEAT      490    519       ANK 6.
FT   REPEAT      542    572       ANK 7.
FT   REPEAT      588    617       ANK 8.
FT   REPEAT      621    650       ANK 9.
FT   REPEAT      654    683       ANK 10.
FT   REPEAT      687    716       ANK 11.
FT   REPEAT      724    753       ANK 12.
FT   REPEAT      769    798       ANK 13.
FT   REPEAT      802    832       ANK 14.
FT   REPEAT      836    865       ANK 15.
FT   REPEAT      870    899       ANK 16.
FT   REPEAT      905    934       ANK 17.
FT   REPEAT      938    967       ANK 18.
FT   REPEAT      971   1001       ANK 19.
FT   REPEAT     1005   1037       ANK 20.
FT   REPEAT     1043   1072       ANK 21.
FT   ZN_FING    1104   1164       FYVE-TYPE.
FT   VARSPLIC    392    439       LDLELKDHEGSTALWLAVQYITVSSDQSVNPFEDLPVVNGT
FT                                SFDENSL -> YVGRPGEMQSGCEGGIIRFRAERSRRQHCT
FT                                LAGRPVHHCVFRPVCKPL (in isoform 2).
FT                                /FTId=VSP_007917.
FT   VARSPLIC    440   1169       Missing (in isoform 2).
FT                                /FTId=VSP_007918.
FT   CONFLICT    439    439       L -> F (IN REF. 1; BAC67389).
FT   CONFLICT    448    448       S -> T (IN REF. 1; BAC67389).
FT   CONFLICT    502    502       G -> V (IN REF. 1; BAC67389).
FT   CONFLICT    527    527       L -> V (IN REF. 1; BAC67389).
FT   CONFLICT    840    840       P -> S (IN REF. 1; BAC67389).
SQ   SEQUENCE   1169 AA;  128603 MW;  A78E5406A0AD0165 CRC64;
```

```
  Query Match             7.0%;  Score 453.5;  DB 1;  Length 1169;
  Best Local Similarity   24.2%;  Pred. No. 7.8e-19;
  Matches  201;  Conservative  125;  Mismatches  311;  Indels  195;  Gaps   25;

Qy        102 LFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQAR 161
              |  :  : ||:     |:    : :  ||| : ||||  | :  |       :  |:| :
Db        293 LHKGIQRGDLFASTFLIKNGALVNAATAGAQETPLHLVALYSPK-------KYSADVMSE 345

Qy        162 DDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQ-H 220
                              |::   ||: ||:|| :|:   |||| : :      |   ||| |
Db        346 ---------------MAQIAEALLQAGANPNMQDSKGRTPLHLSIMARNDCVFSQLLQCK 390

Qy        221 GAEPTIRNTDGRTALDLA--------DPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 272
                :  ::: :| ||| ||        | |          :|  : :  | :|  : | |
Db        391 QLDLELKDHEGSTALWLAVQYITVSSDQSVNPF------EDLPVVNGTSFDENSLAARLI 444
```

```
Qy        273 PLNVNCHASDGRK-STPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHY 331
Db        445 QRGSNTDAPDVMTGNCLLQRAAGAGNEAAALFLATSGAHANHRNKWGETPLHTACRHGLA 504

Qy        332 EVTELLVKHGACVN------------------AMDLWQFTPLHEAASKNRVEVCSLLLS 372
Db        505 NLTAELLQQGANPNLQTEEALTLPKESPVLMSSADSIYLQTPLHMAIAYNHPDVVSVILE 564

Qy        373 YGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNF 432
Db        565 QKANA----LHATNNLQIIPDFSLKD------------------SRDQTVLGLAL--- 597

Qy        433 KHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEV 492
Db        598 ------WTGMHTIAAQ----------LLGSGASINDTMSDGQTLLHMAIQRQDSKSALF 640

Qy        493 VVKHEAKVNALDNLGQTSLHRA----------AYCGHLQTCRLLLSYGCDPNIISLQG-- 540
Db        641 LLEHQADINVRTQDGETALQLAIKHQLPLVVDAIC----------TRGADMSVPDEKGNP 690

Qy        541 --FTALQMGNENV-QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 597
Db        691 PLWLALASNLEDIASTLVRHG-------CDATCWGPGPSGCLQTL--------------- 728

Qy        598 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHA------------KDKGGLVPLHNACSYGH 645
Db        729 ------LHRAVDENNESTACFLIRSGCDVNSPRQPGTNGEGEEEARDGQTPLHLAASWGL 782

Qy        646 YEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH-GADPTKKNRDGNTPLD 704
Db        783 EETVQCLLEFGANVNAQDAEGRTPVHVAISNQHSVIIQLLISHPNIELSVRDRQGLTPF- 841

Qy        705 LVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLS-SPDNVNCRDTQGRHSTPLHLAAGY 763
Db        842 --------------ACAMTYKNNKAAEAILKRESGAAEQV---DNKGRNF--LHVAVQN 881

Qy        764 NNLEVAEYLLQHGADVNA--QDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFT 821
Db        882 SDIESVLFLISVQANVNSRVQDASKLTPLHLAVQAGSEIIVRNLLLAGAKVNELTKHRQT 941

Qy        822 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLV----SADDVSALLT 869
Db        942 ALHLAAQQDLPTICSVLLENGVDFAAVDENGNNALHLAVMHGRLNNIRALLT 993
```

```
RESULT 11
AFY1_HUMAN
ID   AFY1_HUMAN     STANDARD;      PRT;  1169 AA.
AC   Q9P2R3; Q9ULG5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeats- and FYVE domain-containing protein 1 (Ankyrin repeats
DE   hooked to a zinc finger motif).
GN   ANKFY1 OR ANKHZN OR KIAA1255.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC   TISSUE=Fetal brain;
RX   MEDLINE=20400338; PubMed=10940552;
RA   Kuriyama H., Asakawa S., Minoshima S., Maruyama H., Ishii N., Ito K.,
RA   Gejyo F., Arakawa M., Shimizu N., Kuwano R.;
RT   "Characterization and chromosomal mapping of a novel human gene,
RT   ANKHZN.";
RL   Gene 253:151-160(2000).
RN   [2]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Brain;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K.-I., Kikuno R., Hirosawa M., Nomura N.,
RA   Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Also associated with endosomes.
CC   -!- TISSUE SPECIFICITY: High expression in whole adult brain and
CC       intermediate expression in all other tissues and specific brain
CC       regions examined, including fetal brain.
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC   -!- SIMILARITY: Contains 21 ANK repeats.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB037360; BAA90300.1; -.
DR   EMBL; AB033081; BAA86569.2; ALT_INIT.
DR   HSSP; P42773; 1BU9.
DR   Genew; HGNC:20763; ANKFY1.
DR   MIM; 607927; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF00023; ank; 20.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF01363; FYVE; 1.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00225; BTB; 1.
DR   SMART; SM00064; FYVE; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 12.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
KW   ANK repeat; Repeat; Zinc-finger.
FT   DOMAIN       68    130       BTB.
FT   REPEAT      217    247       ANK 1.
FT   REPEAT      255    284       ANK 2.
FT   REPEAT      288    317       ANK 3.
FT   REPEAT      322    362       ANK 4.
FT   REPEAT      366    396       ANK 5.
FT   REPEAT      490    519       ANK 6.
FT   REPEAT      542    572       ANK 7.
FT   REPEAT      588    617       ANK 8.
FT   REPEAT      621    650       ANK 9.
FT   REPEAT      654    683       ANK 10.
FT   REPEAT      687    716       ANK 11.
FT   REPEAT      724    753       ANK 12.
FT   REPEAT      769    798       ANK 13.
FT   REPEAT      802    832       ANK 14.
FT   REPEAT      836    865       ANK 15.
FT   REPEAT      870    899       ANK 16.
FT   REPEAT      905    934       ANK 17.
FT   REPEAT      938    967       ANK 18.
FT   REPEAT      971   1001       ANK 19.
FT   REPEAT     1005   1037       ANK 20.
FT   REPEAT     1043   1072       ANK 21.
FT   ZN_FING    1104   1164       FYVE-TYPE.
FT   CONFLICT      1      4       MAEE -> AGGK (IN REF. 2).
FT   CONFLICT     48     48       E -> K (IN REF. 1).
FT   CONFLICT    115    115       M -> V (IN REF. 1).
FT   CONFLICT    207    207       A -> P (IN REF. 1).
FT   CONFLICT    235    235       F -> S (IN REF. 1).
FT   CONFLICT    270    270       S -> N (IN REF. 1).
FT   CONFLICT    283    291       DMVDKSGWS -> AWWPRVLE (IN REF. 1).
FT   CONFLICT    299    300       RG -> E (IN REF. 1).
FT   CONFLICT    322    322       A -> C (IN REF. 1).
FT   CONFLICT    336    337       KK -> RN (IN REF. 1).
```

```
FT   CONFLICT    379    379       N -> D (IN REF. 1).
FT   CONFLICT    385    385       Q -> H (IN REF. 1).
FT   CONFLICT    645    645       Q -> R (IN REF. 1).
FT   CONFLICT    652    654       TQD -> PQA (IN REF. 1).
FT   CONFLICT    687    687       K -> E (IN REF. 1).
FT   CONFLICT    713    713       D -> G (IN REF. 1).
FT   CONFLICT    784    784       MISSING (IN REF. 1).
FT   CONFLICT    795    795       N -> D (IN REF. 1).
FT   CONFLICT    798    798       A -> P (IN REF. 1).
FT   CONFLICT    807    807       I -> C (IN REF. 1).
FT   CONFLICT    827    829       DIH -> ISS (IN REF. 1).
FT   CONFLICT    835    835       R -> K (IN REF. 1).
FT   CONFLICT    849    849       N -> D (IN REF. 1).
FT   CONFLICT    860    860       E -> G (IN REF. 1).
FT   CONFLICT    874    874       F -> S (IN REF. 1).
FT   CONFLICT    889    889       F -> S (IN REF. 1).
FT   CONFLICT    901    901       V -> A (IN REF. 1).
FT   CONFLICT    905    905       S -> P (IN REF. 1).
FT   CONFLICT    913    913       A -> V (IN REF. 1).
FT   CONFLICT    916    916       A -> E (IN REF. 1).
FT   CONFLICT    933    933       N -> T (IN REF. 1).
FT   CONFLICT    940    940       Q -> K (IN REF. 1).
FT   CONFLICT    949    949       Q -> E (IN REF. 1).
FT   CONFLICT    962    962       G -> A (IN REF. 1).
FT   CONFLICT   1120   1120       T -> A (IN REF. 1).
FT   CONFLICT   1157   1157       N -> T (IN REF. 1).
SQ   SEQUENCE   1169 AA;  128398 MW;  589297CA4ACDDB56 CRC64;

  Query Match             6.9%;  Score 448;  DB 1;  Length 1169;
  Best Local Similarity   24.4%;  Pred. No. 1.6e-18;
  Matches  205;  Conservative   99;  Mismatches  355;  Indels  180;  Gaps   23;

Qy        120 PEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAE 179
              | |:|  |  |     |  |     : :   |: : |:|   |  |   ||     |
Db        246 PGKLNEADHNG--DLALDLALSRRLESIATTLVSHKADVDMVDKSGWSLLHKGIQRGDLF 303

Qy        180 VVNLLLRHGADPNARD-NWNYTPLHEAAI----KGKIDV-------CIVLLQHGAEPTIR 227
                |:::||  ||       ||||  |:    |   ||          ||| || | ::
Db        304 AATFLIKNGAFVNAATLGAQETPLHLVALYSSKKHSADVMSEMAQIAEALLQAGANPNMQ 363

Qy        228 NTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKST 287
              :: ||| | :                 |  :|||    ||   ::     |   ||
Db        364 DSKGRTPLHV-------------------SIMAGNEYVFSQLLQCKQLDLELKDHEGST 403

Qy        288 PLHLA----------------------AGYNRVKIVQLLLQHGADVHAKD--KGDLVPL 322
              | ||                      ::       |:| |:   | |   |: : |
Db        404 ALWLAVQHITVSSDQSVNPFEDVPVVNGTSFDENSFAARLIQRGSHTDAPDTATGNCL-L 462

Qy        323 HNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNC 382
                |   |:      |  :|| ||  : |  |||| |       : : ||  ||:| |
Db        463 QRAAGAGNEAAALFLATNGAHVNHRNKWGETPLHTACRHGLANLTAELLQQGANPNL--- 519

Qy        383 HNKSAIDLAPTPQLKERLAYEPKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETAL 442
                  | :| |    :  ||   |           ||          :| |
Db        520 -----------QTEEALPLPKEAASLTSLADSV-------HL------------QTPL 547

Qy        443 HCAAASPYPKRKQICELLLRKGANINEKTKEF-------------LTPLHVASEKAHNDV 489
              | | |  :|    :  ::| : ||    |               | | :|    : :
Db        548 HMAIAYNHP---DVVSVILEQKANALHATNNLQIIPDFSLKDSRDQTVLGLALWTGMHTI 604

Qy        490 VEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNE 549
              ::    | :|   : ||| || |       ::   || :  | |: :  | ||||:
Db        605 AAQLLGSGAAINDTMSDGQTLLHMAIQRQDSKSALFLLEHQADINVRTQDGETALQLAIR 664

Qy        550 NVQQLLQEGISLGNS------EADRQLLEAAKAGDVETVKKLCTVQSVNCRDI------E 597
              |   |: : |    :      |     | | | ::| :    |:     |
Db        665 NQLPLVVDAICTRGADMSVPDEKGNPPLWLALANNLEDIAS--TLVRHGCDATCWGPGPG 722

Qy        598 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHA-----------KDKGGLVPLHNACSYGH 645
              |   | || |   |       :|:: | ||::           : : |  ||| | |:|
Db        723 GCLQTLLHRAIDENNEPTACFLIRSGCDVNSPRQPGANGEGEEEARDGQTPLHLAASWGL 782

Qy        646 YEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH-GADPTKKNRDGNTPLD 704
              |   | || |   |       :|:: | ||::           : : |  ||| | |:|
Db        783 EETVQCLLEFGANVNAQDAEGRTPIHVAISSQHGVIIQLLVSHPDIHLNVRDRQGLTPF- 841

Qy        705 LVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCR--------DTQGRHSTP 756
                                              |   |   | : :  |        | :||:
Db        842 ------------------------ACAMTFKNNKSAEAILKRESGAAEQVDNKGRNF-- 874

Qy        757 LHLAAGYNNLEVAEYLLQHGADVNA--QDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
              ||:|   :::|   :|:   |:||:  ||   | ||| |   |   :   |:   | ||
Db        875 LHVAVQNSDIESVLFLISVHANVNSRVQDASKLTPLHLAVQAGSEIIVRNLLLAGAKVNE 934

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLV----SADDVSALLT 869
                |   | || |||:    :|::|| :| |   :: |   | |       :::  |||
Db        935 LTKHRQTALHLAAQQDLPTICSVLLENGVDFAAVDENGNNALHLAVMHGRLNNIRVLLT 993


RESULT 12
V162_FOWPV
ID   V162_FOWPV     STANDARD;      PRT;   603 AA.
AC   Q9J569;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV162.
GN   FPV162.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 14 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF198100; AAF44506.1; -.
DR   HSSP; P25963; 1IKN.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       23     53       ANK 1.
FT   REPEAT       57     87       ANK 2.
FT   REPEAT       91    120       ANK 3.
FT   REPEAT      124    155       ANK 4.
FT   REPEAT      159    189       ANK 5.
FT   REPEAT      193    223       ANK 6.
FT   REPEAT      227    257       ANK 7.
FT   REPEAT      261    291       ANK 8.
FT   REPEAT      295    325       ANK 9.
FT   REPEAT      329    362       ANK 10.
FT   REPEAT      366    397       ANK 11.
FT   REPEAT      401    428       ANK 12.
FT   REPEAT      432    467       ANK 13.
FT   REPEAT      504    533       ANK 14.
SQ   SEQUENCE   603 AA;  66542 MW;  8FFE3E464DB9D16D CRC64;

  Query Match             6.7%;  Score 431;  DB 1;  Length 603;
  Best Local Similarity   22.9%;  Pred. No. 6.2e-18;
  Matches  168;  Conservative  103;  Mismatches  234;  Indels  230;  Gaps   23;
```

```
Qy        132 KSTPLHFAAGFGRKDVVEYL---LQNGANVQARDDGGLIPLHNAC-SFGHAEVVNLLLRH 187
Db         24 KDTPLHKAVML--PDAVERIRMFVSKGADINVISDFKKTALHYAAKKLATPEVLKTLIYL 81

Qy        188 GADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLT 247
Db         82 GTNVNVTDMFESTPLHYAVQENGLEATKKLLDLGADPNTKYMNGQTPLHCA---AMVIPD 138

Qy        248 GEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY-NRVKIVQLLLQ 306
Db        139 G---------------PELVRILVEYGANVNALDNKHNTPLALAAELSNTNKTIETLIE 182

Qy        307 HGADVHAKDKGDLVPLHNAC-SYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS----K 361
Db        183 LGADVKIKNNDGITPLHLAAKSSSDSKTVETLILHGADVNATCSEGNTPLHDAATSYELS 242

Qy        362 NRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRI 421
Db        243 NTIE---MLIEYGAEVNAAN---------------------------------------- 259

Qy        422 KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVA 481
Db        260 -------------SVGDTPLHCAARSRNPVHK--LKTLIAHGSNVNAVNGISVTPLHLA 303

Qy        482 --SEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQ 539
Db        304 TYSDNA-TEALKVLIEHGAEVNSVDIYGRTPMH------------------------- 335

Qy        540 GFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR 599
Db        336 ----------------------------------------YISRSYSSQSLK------- 347

Qy        600 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE--VAELLVKHGA 657
Db        348 -------------TAVELLVEHGADIEAKNVIGGTPLSSACNNIEYDLRLIECFIEYGA 393

Qy        658 VVNVADLWKFTPLHEAAAKGKY-EICKLLLQHGADPTKKNRDGNTPLD-----LVKDGDT 711
Db        394 DINTRDIRDETPLYSAI---KYPEIVNLLMNYSASTNITNKSNITPLESAIANCIGSAEI 450

Qy        712 DIQDL---------LRGDAALLDAAK--KGC--LARVKKLSSPDNVNCRDTQGRHSTPLH 758
Db        451 IVTQIILDAFRFPDIKNDAIFIRNMKTIEECTMLIDVKESCEYDINKMRSIKFNNMYGLD 510

Qy        759 LAAGYNNLEVAEYLLQHGADVNAQD----------------------KGGLIPLHNAA 794
Db        511 IFIRSNNINLLSSLVSNVEDIYLEPGCFLVYGNKLRKSVYAARKRLSLLKNSISILSNIT 570

Qy        795 SYGHVDVAALLIKYN 809
Db        571 TDGYWNALPIELKYN 585
```

```
RESULT 13
V222_FOWPV
ID   V222_FOWPV     STANDARD;      PRT;   747 AA.
AC   Q9J513;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative ankyrin-repeat protein FPV222.
GN   FPV222.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 14 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF198100; AAF44566.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001810; F-box.
DR   Pfam; PF00023; ank; 12.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       38     67       ANK 1.
FT   REPEAT      103    132       ANK 2.
FT   REPEAT      136    165       ANK 3.
FT   REPEAT      169    198       ANK 4.
FT   REPEAT      202    231       ANK 5.
FT   REPEAT      234    263       ANK 6.
FT   REPEAT      294    323       ANK 7.
FT   REPEAT      328    357       ANK 8.
FT   REPEAT      361    393       ANK 9.
FT   REPEAT      397    426       ANK 10.
FT   REPEAT      430    460       ANK 11.
FT   REPEAT      464    493       ANK 12.
FT   REPEAT      495    524       ANK 13.
FT   REPEAT      529    559       ANK 14.
SQ   SEQUENCE   747 AA;  85303 MW;  55F90AF2855C3D28 CRC64;
```

Query Match 6.2%; Score 399.5; DB 1; Length 747;
Best Local Similarity 22.7%; Pred. No. 6.1e-16;
Matches 184; Conservative 126; Mismatches 306; Indels 193; Gaps 28;

```
Qy        144 RKDVVEY--LLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
Db         16 KNDVVSVRKYLEKGINPNEKNKDNCTMLYTAVEHRYIDIIKLLLDHGADPNIYSSDHMTP 75

Qy        202 LHEAAIKGKI-DVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
Db         76 LHSVSVIIPIRKISKLITKYGNLVTLANYRNTFFV-------------YNENRNLEIA- 120

Qy        261 SGNEEKMM----ALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDK 316
Db        121 -----KMLIQNGALVNMNNM-------KNITPLHIASSSGSYKMVELLLLHGANTNALTS 168

Qy        317 GDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGAD 376
Db        169 YGETSLHYSVSSNDLNISELLIENGTNVNVANKDSITALIIAVEIMSIDLVRLLLDKGAD 228

Qy        377 PTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFK--- 433
Db        229 TNAIGL-ERFKLYVTETKQNNNILKY------LNTNNVNTNVTMINEYIASELYDWNRNS 281

Qy        434 ----------HPQTHETALHCAAASPYPKRKQICELLLRKGAN--INEKTKEFLTPLHV 480
Db        282 ATSKLMFRSCFEPCTVPVTLATRKGS-----KELLEILLEYGCNPDICEKTTSTYA-MHY 335

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
Db        336 AVIRKHYEMLNILIRYDAYTDVKDRQQNTPAHYAVKLPISESCKYL------------- 381

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        382 ---------------------------KLLKLAGA------------SFNLTNRKGR- 399

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAV-V 659
```

```
Db        400 -TPLHTACKYNNTEAVKYLIESGCDTNIVDVMSFTPLNYAVYYEREDTVKILLESGCVDP 458

Qy        660 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNR-DGNTPLDLVKDGDTDIQDLLR 718

Db        459 NLCDYKEVSPIIQAIKRNNKNIIKMLLNAGIDIKPINECYGLHMLAALHN-----KDLL- 512

Qy        719 GDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPL-HLAAGYNNLEVAEYLLQHGA 777

Db        513 ----------KWLLCTISEL----EVNGVDD---HYVPLASYVAELSDIRIMEILIEKGL 555

Qy        778 DVN---AQDKGGLIPLHNAASYGHVDVAALLI-------------KYNACVNATDKWAFT 821

Db        556 DLNKVTGPDETMFTMIFSATSDLRKSIIDLLISQIAADEEFSEGFKINKNMIQTDKYLLR 615

Qy        822 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS---ADDVSALLTAAMPPSAL- 877

Db        616 VYHEC----KNQVSKM-------GEIKLGDGFTMIDIYKNRRSIHVNFLARYAMQLSTID 664

Qy        878 ----------------PSCYKPQVLNGVR 890

Db        665 LREVPIYRKYLEILINPAIKRHKILNAAK 693


RESULT 14
ANR3_HUMAN
ID   ANR3_HUMAN     STANDARD;      PRT;   832 AA.
AC   P57078; Q96KH0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat
DE   domain protein 3) (PKC-delta-interacting protein kinase).
GN   ANKRD3 OR DIK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Fetal kidney, and Fetal lung;
RA   Shimizu N., Kudoh J., Shibuya K.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P57078-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P57078-2; Sequence=VSP_004862;
CC   -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB047783; BAB56136.1; -.
DR   EMBL; AP001743; BAA95526.1; -.
DR   HSSP; P25963; 1IKN.
DR   Genew; HGNC:496; ANKRD3.
DR   MIM; 605706; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00023; ank; 10.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   ANK repeat; Alternative splicing.
FT   DOMAIN       22    286       PROTEIN KINASE.
FT   REPEAT      485    514       ANK 1.
FT   REPEAT      518    547       ANK 2.
FT   REPEAT      551    580       ANK 3.
FT   REPEAT      584    613       ANK 4.
FT   REPEAT      617    647       ANK 5.
FT   REPEAT      651    680       ANK 6.
FT   REPEAT      684    713       ANK 7.
FT   REPEAT      717    746       ANK 8.
FT   REPEAT      750    780       ANK 9.
FT   REPEAT      782    811       ANK 10.
FT   NP_BIND      28     36       ATP (BY SIMILARITY).
FT   BINDING      51     51       ATP (BY SIMILARITY).
FT   ACT_SITE    143    143       BY SIMILARITY.
FT   VARSPLIC    278    325       Missing (in isoform 2).
FT                                /FTId=VSP_004862.
FT   CONFLICT    714    714       M -> V (IN REF. 1).
SQ   SEQUENCE   832 AA;  91610 MW;  5D8FFFD5F04F7ECB CRC64;

  Query Match             6.0%;  Score 390.5;  DB 1;  Length 832;
  Best Local Similarity   25.3%;  Pred. No. 2.4e-15;
  Matches  142;  Conservative   75;  Mismatches  186;  Indels  158;  Gaps   14;

Qy        178 AEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNT-------- 229

Db        359 SEVVPARLKRASAPTFDNDYSLSEL-----LSQLDSGVSQAVEGPEELSRSSSESKLPSS 413

Qy        230 ------DGRTALDLA------------DPSAKAVLTGEYKKDELLESARSGNEEKMMAL 270

Db        414 GSGKRLSGVSSVDSAFSSRGSLSLSFEREPSTSDLGTTDVQKKKLVDAIVSGDTSKLMKI 473

Qy        271 LTPL-----------------------------NVNCHASDGRKSTPLHLAAGYNRVK 299

Db        474 LQPQDVDLALDSGASLLHLAVEAGQEECAKWLLLNNANPNLSNRRGSTPLHMAVERRVRG 533

Qy        300 IVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAA 359

Db        534 VVELLLARKISVNAKDEDQWTALHFAAQNGDESSTRLLLEKNASVNEVDFEGRTPMHVAC 593

Qy        360 SKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVT 419

Db        594 QHGQENIVRILLRRGVDVSL---QGKDA------------------------------- 618

Qy        420 RIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK-GANINEKTKEFLTPL 478

Db        619 -------------------WLPLHYAA---WQGHLPIVKLLAKQPGVSVNAQTLDGRTPL 656
```

```
Qy        479 HVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISL 538
Db        657 HLAAQRGHYRVARILIDLCSDVNVCSLLAQTPLHVAAETGHTSTARLLLHRGAGKEAMTS 716

Qy        539 QGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEG 598
Db        717 DGYTALHL------------------------AARNGHLATVKLLVEEKA----DVLA 746

Qy        599 R---QSTPLHFAAGYNRVSVVEYLLQHGADV-HAKDKGGLVPLHNACSYGHYEVAELLVK 654
Db        747 RGPLNQTALHLAAAHGHSEVVEELV--SADVIDLFDEQGLSALHLAAQGRHAQTVETLLR 804

Qy        655 HGAVVNVADLWKFTPLHEAAA 675
Db        805 HGAHINLQSL-KFQGGHGPAA 824
```

```
RESULT 15
V024_FOWPV
ID   V024_FOWPV     STANDARD;      PRT;   596 AA.
AC   Q9J5H7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV024.
GN   FPV024.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 13 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF198100; AAF44368.1; -.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT        5     34       ANK 1.
FT   REPEAT       37     66       ANK 2.
FT   REPEAT       68     96       ANK 3.
FT   REPEAT       99    128       ANK 4.
FT   REPEAT      130    158       ANK 5.
FT   REPEAT      162    191       ANK 6.
FT   REPEAT      195    225       ANK 7.
FT   REPEAT      229    258       ANK 8.
FT   REPEAT      262    291       ANK 9.
FT   REPEAT      295    324       ANK 10.
FT   REPEAT      326    355       ANK 11.
FT   REPEAT      359    389       ANK 12.
FT   REPEAT      394    423       ANK 13.
SQ   SEQUENCE   596 AA;  67938 MW;  F813C43228907D07 CRC64;

  Query Match             5.9%;  Score 382.5;  DB 1;  Length 596;
  Best Local Similarity   21.7%;  Pred. No. 4.4e-15;
  Matches  156;  Conservative  103;  Mismatches  218;  Indels  241;  Gaps   21;
```

```
Qy         96 EPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNG 155
Db          4 EKLRKDLHRTIRTRDLNAVKHIILKKYTFSNKNA--LSTPLYLAVSNSDIDIVKFLLDNG 61

Qy        156 ANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCI 215
Db         62 ADINKCKSP---PLHKAINLGNVEMVKLLVDHGADIEKVYLGN-SPLYLALCKRNTNITK 117

Qy        216 VLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
Db        118 YLLERGADP---NT---LFINYCDAIYNKIPIDIFK-----------------ILIKYK 153

Qy        276 VNCHASDGRKSTPLHLA----------------------AGYNR------------VKIV 301
Db        154 VSLNIQNSHFKTPIYYAIKCTNYPLIKLLLENNASLTIPEGYNNHYLITAVKHNCDISIL 213

Qy        302 QLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASK 361
Db        214 RLLIKYGVPVNEQDDLERTSLHYCVSAGKHDILKLLLDYDADPNITDSCLGTPLHYAVSR 273

Qy        362 NRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRI 421
Db        274 NDIIATTLLIEKGAN---VNIHN---------------------------------------- 293

Qy        422 KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVA 481
Db        294 -------------DTIDTVLNIAVGN---RNKILINLLLMYGANTRLKSRNPL--IHKA 334

Qy        482 SEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC-GHLQTCRLLLSYGCDPNIISLQG 540
Db        335 LETKDINILSEILNHGAEVNIYNREGYTPLYIAIITFMQIKFAKLLLRYGSNPN------ 388

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        389 -----MKNES---------------------------------------------------NE 395

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        396 NTPLHGAILSNRLDSVELLMSYNVDVHSINKLGHTPL-SCINYISDKIATII-------- 446

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
Db        447 -------------------ISKIV----------------LDLEKDSNLFLLDGFKAN 469

Qy        721 AALLD------AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYL 772
Db        470 IECIDQNDRFKVIRKNCEDELKSIR---NIKLNH---RYSLSIFLHSDNNNNILIRFL 521
```

```
RESULT 16
V244_FOWPV
ID   V244_FOWPV     STANDARD;      PRT;   668 AA.
AC   Q9J4Z6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV244.
GN   FPV244.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 12 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44588.1; -.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 10.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       40     69       ANK 1.
FT   REPEAT      144    173       ANK 2.
FT   REPEAT      177    206       ANK 3.
FT   REPEAT      210    239       ANK 4.
FT   REPEAT      272    302       ANK 5.
FT   REPEAT      306    336       ANK 6.
FT   REPEAT      340    370       ANK 7.
FT   REPEAT      374    403       ANK 8.
FT   REPEAT      407    437       ANK 9.
FT   REPEAT      441    471       ANK 10.
FT   REPEAT      473    502       ANK 11.
FT   REPEAT      571    602       ANK 12.
SQ   SEQUENCE   668 AA;  76200 MW;  F3C5BA076023791B CRC64;

  Query Match             5.9%;  Score 380.5;  DB 1;  Length 668;
  Best Local Similarity   27.3%;  Pred. No. 6.8e-15;
  Matches  153;  Conservative   95;  Mismatches  192;  Indels  121;  Gaps   26;

Qy        326 CSYGHYEVTELLV-----KHGA--CVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPT 378
              |   : |:   :       ||||  | |  :   || :|:|    ::::   |:    :|
Db         11 CRGNYIEILSAITNYNLHKHGANQCEN--ESIPFTAIHQALQLRQIDIVKELIQ--QNPK 66

Qy        379 LL----NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
              |:    :  | :   :  || : :          ::    : |:    |: |: ::| ||
Db         67 LIYVTDHRRNSTLHTICITPNVMD---------IVISLTVDCDIILDIKYASI-ILN-KH 115

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINE----KTKEFLTPLHVASEKAHND-- 488
                                |  : |  :|::| : ||    |  : :  : :  |:   |
Db        116 ----------------KLGEACIHVLKEGISGNEISYNKINKSIEYMKLIKERIQQDEL 158

Qy        489 -VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ-- 545
               : |:::|    ||| |   :| :| ||   |: :   ||||||  | |||:|   : |:
Db        159 LIAEMLLKKGIDVNAKDVYCRTPIHYAAERGNTKMVNLLLSYGADVNIITLDDLSVLEYA 218

Qy        546 MGNENVQQLLQEGISLGN-SEADRQLLEAAKAGDVETVKKLC-TVQSVNCRDIEGRQSTP 603
              : ::|:  :     :  | :: |  ||:| :  |:||   |  :  |||  |:   ::||
Db        219 VDSKNIDTIKAIIDNRSNINKNDLSLLKAIRNTDLETSLLLYDSGFSVNSIDV--YKNTP 276

Qy        604 LHFAAGYNRVS-VVEYLLQHGADVHAKDKGGLVPLHNACSYGH-YEVAELLVKHGAVVNV 661
              ||:|     :| :|  ||: | ||:||:  |  ||:     |:  |     |:  || ||
Db        277 LHYAVQAPSLSRLVPKLLERGIDVNAKNIKGETPLYLMAKNGYDTENIRTLIMRGADVNA 336

Qy        662 ADLWKFTPLHEAAAKGKYEICKL-LLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
              ||    ||||:|:   :|:   : ||: ||
Db        337 ADSLYITPLHQASTLDRYKDTVITLLELGA---------------------------- 366

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
                                  ||| ||   :  ||:| ||  ||: :    || :|||:
Db        367 ---------------------NVNARDYCDK--TPIHYAAVRNNVVIINTLLDYGADIE 402

Qy        781 A-QDKGGLIPLHNAASYG---HVDVAALLIKYNACVNATDKWAFTPLHEAAQKG-RTQLC 835
              |   | | : ||  | ||   :: |    ||   | ||: :|:  |||| | :|  : ::
Db        403 ALSQKIGTV-LH-FALYGTNPYMSVKT-LIDRGANVNSKNKYLSTPLHYACKKNCKPEVI 459

Qy        836 ALLLAHGADPTLKNQEGQTPL 856
                :|| :|||    |   | ||
Db        460 KMLLDNGADVNAINIRNQYPL 480

RESULT 17
```

```
DAK1_HUMAN
ID   DAK1_HUMAN     STANDARD;      PRT;  1432 AA.
AC   P53355; Q9BTL8;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Death-associated protein kinase 1 (EC 2.7.1.37) (DAP kinase 1).
GN   DAPK1 OR DAPK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, INDUCTION, AND MUTAGENESIS OF LYS-42.
RX   MEDLINE=95129831; PubMed=7828849;
RA   Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT   "Identification of a novel serine/threonine kinase and a novel 15-kD
RT   protein as potential mediators of the gamma interferon-induced cell
RT   death.";
RL   Genes Dev. 9:15-30(1995).
RN   [2]
RP   REVISIONS TO 164-171.
RA   Feinstein E.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-367 FROM N.A.
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RT   "Cloning of human full-length CDSs in BD Creator(TM) system donor
RT   vector.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=20094983; PubMed=10629061;
RA   Inbal B., Shani G., Cohen O., Kissil J.L., Kimchi A.;
RT   "Death-associated protein kinase-related protein 1, a novel
RT   serine/threonine kinase involved in apoptosis.";
RL   Mol. Cell. Biol. 20:1044-1054(2000).
RN   [5]
RP   FUNCTION, ENZYME REGULATION, MUTAGENESIS OF LYS-42; SER-308 AND
RP   SER-313, AND PHOSPHORYLATION OF SER-308.
RX   PubMed=11579085;
RA   Shohat G., Spivak-Kroizman T., Cohen O., Bialik S., Shani G.,
RA   Berissi H., Eisenstein M., Kimchi A.;
RT   "The pro-apoptotic function of death-associated protein kinase is
RT   controlled by a unique inhibitory autophosphorylation-based
RT   mechanism.";
RL   J. Biol. Chem. 276:47460-47467(2001).
CC   -!- FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase
CC       which acts as a positive regulator of apoptosis.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- COFACTOR: Magnesium.
CC   -!- ENZYME REGULATION: Negatively regulated by autophosphorylation on
CC       Ser-308.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Colocalizes with the actin
CC       filament system.
CC   -!- INDUCTION: Up-regulated following treatment with interferon-gamma.
CC   -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
CC       kinase subfamily.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to
CC       frameshifts in positions 466 and 477.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X76104; CAA53712.1; ALT_FRAME.
```

```
DR   EMBL; BT006935; AAP35581.1; ALT_TERM.
DR   PDB; 1JKT; 01-APR-02.
DR   Genew; HGNC:2674; DAPK1.
DR   MIM; 600831; -.
DR   GO; GO:0004684; F:calmodulin-dependent protein kinase I activity; TAS.
DR   GO; GO:0008624; P:induction of apoptosis by extracellular sig. . .; TAS.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00023; ank; 8.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 7.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW   Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW   3D-structure.
FT   DOMAIN       13    275       PROTEIN KINASE.
FT   DOMAIN      267    334       CALMODULIN-BINDING.
FT   REPEAT      378    407       ANK 1.
FT   REPEAT      411    440       ANK 2.
FT   REPEAT      444    475       ANK 3.
FT   REPEAT      479    508       ANK 4.
FT   REPEAT      512    541       ANK 5.
FT   REPEAT      545    574       ANK 6.
FT   REPEAT      578    607       ANK 7.
FT   REPEAT      611    640       ANK 8.
FT   REPEAT      877    906       ANK 9.
FT   REPEAT     1164   1198       ANK 10.
FT   DOMAIN     1314   1398       DEATH.
FT   NP_BIND      19     27       ATP (BY SIMILARITY).
FT   BINDING      42     42       ATP.
FT   ACT_SITE    139    139       BY SIMILARITY.
FT   MOD_RES     308    308       PHOSPHORYLATION (AUTO-).
FT   MUTAGEN      42     42       K->A: LOSS OF ACTIVITY, APOPTOTIC
FT                                FUNCTION AND OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     308    308       S->A: ELEVATED CA(2+)-CALMODULIN BINDING
FT                                AND CA(2+)-CALMODULIN-INDEPENDENT KINASE
FT                                ACTIVITY. INCREASES APOPTOTIC ACTIVITY.
FT   MUTAGEN     308    308       S->D: REDUCED CA(2+)-CALMODULIN BINDING
FT                                AND CA(2+)-CALMODULIN-INDEPENDENT KINASE
FT                                ACTIVITY. DECREASES APOPTOTIC ACTIVITY.
FT   MUTAGEN     313    313       S->A: MINIMAL EFFECT ON ACTIVITY.
FT   CONFLICT    364    366       LGS -> SKR (IN REF. 3).
SQ   SEQUENCE   1432 AA;  160220 MW;  9B88B66C21CAC16C CRC64;

  Query Match             5.7%;  Score 371;  DB 1;  Length 1432;
  Best Local Similarity  22.2%;  Pred. No. 7.6e-14;
  Matches  238;  Conservative  146;  Mismatches  387;  Indels  302;  Gaps   43;

Qy       132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191
Db        33 KSTGLQYAAKFIKK----RRTKSSRRGVSRED-------------IEREVSILKEIQHP 74

Qy       192 NARDNWNYTPLHEAAIKGKIDVCIVL-LQHGAE--PTIRNTDGRTALDLADPSAKAVLTG 248
Db        75 NV------ITLHE-VYENKTDVILILELVAGGELFDFLAEKESLTE-EEATEFLKQILNG 126

Qy       249 EYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHG 308
Db       127 VYYLHSLQIAHFDLKPENIMLL----------DRNVPKP--------RIKIIDFGLAHK 167

Qy       309 ADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCS 368
Db       168 IDPGNEFKN---------IFGTPEFV----------APEIVNYEPLGLEADMWSIGVIT 207

Qy       369 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH--SLLQAAREADVTRI----- 421
Db       208 YILLSGASPFLGDTKQETLANVS-------AVNYEFEDEYFSNTSALAKDFIRRLLVKDP 260

Qy       422 KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN-EKTKEFLTPLHV 480
Db       261 KKRMTIQ---------DSLQH-----PWIKPKDTQQALSRKASAVNMEKFKKF------ 299

Qy       481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR-------------LL 526
Db       300 AARKKWKQSVRL-------ISLCQRLSRSFLSRS----NMSVARSDDTLDEEDSFVMKAI 348

Qy       527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
Db       349 IHAINDDNVPGLQHLLG-SLSNYDVNQ--------PNKHGTPPLLIAAGCGNIQILQLLI 399

Qy       587 TVQS-VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGH 645
Db       400 KRGSRIDVQDKGG--SNAVYWAARHGHVDTLKFLSENKCPLDVKDKSGEMALHVAARYGH 457

Qy       646 YEVA--ELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
Db       458 ADVAQVQVLCSFGSNPNIQDKEEETPLHCAAWHGYYSVAKALCEAGCNVNIKNREGETP- 516

Qy       704 DLVKDGDTDIQDLLRGDAALLDAAKKG------CLARVKKLSSPDNVNCRDTQGRHSTPL 757
Db       517 -------------------LLTASARGYHDIVECLA-----EHGADLNACDKDGH--IAL 550

Qy       758 HLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDK 817
Db       551 HLAVRRCQMEVIKTLLSQGCFVDYQDRHGNTPLHVACKDGNMPIVVALCEANCNLDISNK 610

Qy       818 WAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADD---VSALL------ 868
Db       611 YGRTPLHLAANNGILDVVRYLCLMGASVEALTTDGKTAEDLARSEQHEHVAGLLARLRKD 670

Qy       869 ------------TAAMPP--------------SALPSCYKPQVLNGV---RSPGATADAL 899
Db       671 THRGLFIQQLRPTQNLQPRIKLKLFGHSGSGKTTLVESLKCGLLRSFFRRRRPRLSSTNS 730

Qy       900 SSGPSSP--SSLSAASSLDNLSGSFSELS----SVVSSSG-TEG---------------- 936
Db       731 SRFPPSPLASKPTVSVSINNLYPGCENVSVRSRSMMFEPGLTKGMLEVFVAPTHHPHCSA 790

Qy       937 ------ASSLEKKEVPGV-DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEI 989
Db       791 DDQSTKAIDIQNAYLNGVGDFSVWEFSGNPVYFCCYDYFAANDPTSIHVVVFSLEEPYEI 850

Qy       990 GIN-------------------AYGHR------------HKLIKGVERLISGQQGL---- 1014
Db       851 QLNPVIFWLSFLKSLVPVEEPIAFGGKLKNPLQVVLVATHADIMNVPRPAGGEFGYDKDT 910

Qy      1015 ----------------NPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV 1050
Db       911 SLLKEIRNRFGNDLHISNKLFVLDAGASG------SKDMKVLRNHLQEIRSQI 957

RESULT 18
ANS1_MOUSE
ID   ANS1_MOUSE     STANDARD;      PRT;  1150 AA.
AC   P59672;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin repeat and SAM domain containing protein 1.
GN   ANKS1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Limb;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 6 ANK repeats.
CC   -!- SIMILARITY: Contains 1 PID domain.
CC   -!- SIMILARITY: Contains 2 SAM domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; BC050847; AAH50847.1; -.
DR   MGD; MGI:2446180; Anks1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR006020; PTB_PID.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 7.
DR   Pfam; PF00640; PID; 1.
DR   Pfam; PF00536; SAM; 2.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00462; PTB; 1.
DR   SMART; SM00454; SAM; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS01179; PID; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 2.
KW   ANK repeat; Repeat.
FT   DOMAIN       31     51       GLY-RICH.
FT   REPEAT       75    104       ANK 1.
FT   REPEAT      108    137       ANK 2.
FT   REPEAT      165    194       ANK 3.
FT   REPEAT      198    227       ANK 4.
FT   REPEAT      231    260       ANK 5.
FT   REPEAT      263    292       ANK 6.
FT   DOMAIN      712    778       SAM 1.
FT   DOMAIN      786    853       SAM 2.
FT   DOMAIN      952   1107       PID.
SQ   SEQUENCE   1150 AA;  125241 MW;  551AF44D1DB02D1B CRC64;

  Query Match             5.7%;  Score 367.5;  DB 1;  Length 1150;
  Best Local Similarity   21.7%;  Pred. No. 8.8e-14;
  Matches  220;  Conservative  135;  Mismatches  330;  Indels  331;  Gaps   37;

Qy        100 RELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQN----- 154
              :|| || | | :  |::|:: ::::|    |   :    :| |      : | :
Db          5 QELLEAARTGHLPAVEKLLSGKRLSSGFGGGGGGSGSGGGSGGGGLGSSSHPLSSLLSMW 64

Qy        155 -GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDV 213
               | ||    |  |  |||:|    || :|| :|||: |  |  |:    ||| || ||  :
Db         65 RGPNVNCVDSTGYTPLHHAALNGHRDVVEVLLRNDALTNVADSKGCYPLHLAAWKGDAQI 124

Qy        214 CIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTP 273
              :|:| |  |:    : : ||::           | ||                      |
Db        125 VRLLIQQG--PSHTRVNEQNALEIR----------ELKK-----------------YGP 154

Qy        274 LNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEV 333
              :  :| :    | || || |   ::|: ||:                          |:
Db        155 FDPYINAKNNDNETALHCAAQYGHTEVVKALLE------------------------EL 189

Qy        334 TELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPT 393
              |: :::    | :   |||  ||   |:||  |||  || | ||:|  :
Db        190 TDPTMRN----NKFE----TPLDLAALYGRLEVVKLLL--GAHPNLLSCSTR-------- 231

Qy        394 PQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKR 453
                                                 ||
Db        232 ------------------------------------KH------------------- 233

Qy        454 KQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHR 513
                                  ||||:|:   |  ||:|::      |    :| ::||
Db        234 ---------------------TPLHLAARNGHKAVVQVLLDAGMDSNYQTEMG-SALHE 270

Qy        514 AAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEA 573
              ||  |    ::||: | | ||   :| |||    : :|: |  :      |:    |:|
Db        271 AALFGKTDVVQILLAAGIDVNIKDNRGLTAL----DTVRDLPSQ-----KSQQIAALIED 321

Qy        574 AKAG--DVETVKKLCTVQSVNCRDIEGRQSTPL--------HFAAGYNRVSVVEYLLQHG 623
                  |  |: | :  | |                    ||          :|    :| | :|
Db        322 HMTGKRSVKEVDRTSTAQ-----------LPLLSNTDAIAPMSQGSMEKTVTELILHF- 368

Qy        624 ADVHAKDKGGLVPLHNACSYGHYE--------------VAELLVKHGAVVNVADL---- 664
              | || ::|    |:|| |   :                ||      | |   :
Db        369 -DTHADEEGPYEALYNAVSCHSLDSTASGRSSDRDSMNKEAEATGTRAAGVRPRERPPPP 427

Qy        665 -------------WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD 710
                          |: ||  ||::|      | ::     :    :   | :|:  :
Db        428 AKPPPDEEEEERVDKKYFPL--AASEG------LAVRPRIQSSAPQEEEEHPYELLLTAE 479

Qy        711 TDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAE 770
              |  : |    |   | : |           | |:|  :| |                 | |
Db        480 T--KKLGTTDGRTEDHRQSG------SGRSQDSVEGQDGQ----------------VPE 514

Qy        771 YL--LQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA- 827
                  |||:   :       |  | |  | : :    : :||          |||:
Db        515 QFSGLLHGSSPVCEVGQDPFQLLTAPSQSHPESS----QQDAC-----------HEASM 558

Qy        828 ------------------QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL 868
                             |  |  | | | |  :  |     ||      |::
Db        559 QLEEPGVQGTEPPQPGVPDQSKRVGLPAGLTALASRTYLDALTHTVPLRPAGAEEED--- 615

Qy        869 TAAMPPSALPSCYKPQV-LNGVRS-PGATADALSSGPSSPSSL-SAASSLDNLS------ 919
              : | |  |   ||:  |   ||   : :| |:  |:  ::: | : || | |
Db        616 -QSGPRSRAPPTSKPKAELKLSRSLSKSDSDLLTCSPTEDATMGSRSESLSNCSIGKKRL 674

Qy        920 -------GSFSELSSVVSSSGTEGASSLEKKEVPG---VDFSITQFVRNLGLEH-----L 964
                    : |:  ::|| |      | |:::: |   :: |: ::: ::||:      |
Db        675 EKSPSFASEWDEIEKIMSSIGEGIDFSQEQQKISGSRTLEQSVGEWLESIGLQQYESKLL 734

Qy        965 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERL----ISGQQGLNP 1016
              :: |:  :     ::|   ::|:|||:    || ||::     |      |  |::|
Db        735 LNGFDDVRFLGSNVME--EQDLREIGISDPQHRRKLLQAARSLPKVKALGYDGVSP 788

RESULT 19
DAK1_MOUSE
ID   DAK1_MOUSE     STANDARD;      PRT;  1442 AA.
AC   Q80YE7; Q80YE6; Q8R341; Q8VDN6; Q9CSD4; Q9JJP7;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
```

```
DE   Death-associated protein kinase 1 (EC 2.7.1.37) (DAP kinase 1).
GN   DAPK1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND TISSUE
RP   SPECIFICITY.
RC   TISSUE=Embryo;
RX   MEDLINE=21523887; PubMed=11485996;
RA   Jin Y., Blue E.K., Dixon S., Hou L., Wysolmerski R.B., Gallagher P.J.;
RT   "Identification of a new form of death-associated protein kinase that
RT   promotes cell survival.";
RL   J. Biol. Chem. 276:39667-39678(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RA   Kimchi A.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6; TISSUE=Brain, and Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 1203-1442 FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase.
CC       Isoform 1 is a negative regulator of apoptosis in direct contrast
CC       to its human homolog. Isoform 2 has no effect on apoptosis and its
CC       function is unknown.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- COFACTOR: Magnesium.
CC   -!- ENZYME REGULATION: Negatively regulated by autophosphorylation on
CC       Ser-308 (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=Beta;
CC         IsoId=Q80YE7-1; Sequence=Displayed;
CC       Name=2; Synonyms=Alpha;
CC         IsoId=Q80YE7-2; Sequence=VSP_050629;
CC   -!- TISSUE SPECIFICITY: High levels in bladder, uterus, vas deferens,
CC       lung, liver and kidney.
CC   -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
CC       kinase subfamily.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AY245540; AAO91934.2; -.
DR   EMBL; AY245541; AAO91935.1; -.
DR   EMBL; X97048; CAA65762.1; -.
DR   EMBL; BC021490; AAH21490.1; ALT_INIT.
DR   EMBL; BC026671; AAH26671.1; -.
DR   EMBL; BC057317; AAH57317.1; -.
DR   EMBL; BC060161; AAH60161.1; -.
DR   EMBL; AK013153; BAB28681.1; -.
DR   MGD; MGI:1916885; Dapk1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00023; ank; 8.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW   Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW   Alternative splicing.
FT   DOMAIN       13    275       PROTEIN KINASE.
FT   DOMAIN      267    334       CALMODULIN-BINDING (BY SIMILARITY).
FT   REPEAT      378    407       ANK 1.
FT   REPEAT      411    440       ANK 2.
FT   REPEAT      444    473       ANK 3.
FT   REPEAT      477    506       ANK 4.
FT   REPEAT      510    539       ANK 5.
FT   REPEAT      543    572       ANK 6.
FT   REPEAT      576    605       ANK 7.
```

```
FT   REPEAT      609    638       ANK 8.
FT   REPEAT      875    904       ANK 9.
FT   REPEAT     1164   1196       ANK 10.
FT   DOMAIN     1312   1396       DEATH.
FT   NP_BIND      19     27       ATP (BY SIMILARITY).
FT   BINDING      42     42       ATP (BY SIMILARITY).
FT   ACT_SITE    139    139       BY SIMILARITY.
FT   MOD_RES     308    308       PHOSPHORYLATION (AUTO-) (BY
FT                                SIMILARITY).
FT   VARSPLIC   1431   1442       Missing (in isoform 2).
FT                                /FTId=VSP_050629.
FT   CONFLICT    354    354       D -> N (IN REF. 2).
FT   CONFLICT    441    441       K -> Q (IN REF. 3;
FT                                AAH57317/AAH60161).
FT   CONFLICT    461    461       V -> A (IN REF. 2).
FT   CONFLICT    960    960       S -> P (IN REF. 2).
FT   CONFLICT   1000   1000       N -> T (IN REF. 2).
FT   CONFLICT   1038   1042       RWLCT -> PMALH (IN REF. 2).
FT   CONFLICT   1105   1105       A -> V (IN REF. 1).
SQ   SEQUENCE   1442 AA;  161440 MW;  243A14D7C6598F63 CRC64;

  Query Match             5.7%;  Score 367;  DB 1;  Length 1442;
  Best Local Similarity   21.7%;  Pred. No. 1.3e-13;
  Matches  232;  Conservative  149;  Mismatches  392;  Indels  296;  Gaps   41;

Qy       132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLL--LRHGA 189
Db        33 KSTGLQYAAKFIKK----RRTKSSRRGVSRED-------------IEREVSILKEIRHP- 74

Qy       190 DPNARDNWNYTPLHEAAIKGKIDVCIVL-LQHGAE--PTIRNTDGRTALDLADPSAKAVL 246
Db        75 --------NVITLHE-VYENKTDVILILELVAGGELFDFLAEKESLTE-EEATEFLKQIL 124

Qy       247 TGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQ 306
Db       125 SGVYYLHSLQIAHFDLKPENIMLL-----------DRNVPKP--------RIKIIDFGLA 165

Qy       307 HGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV 366
Db       166 HKIDFGNEFKN---------IFGTPEFV-----------APEIVNYEPLGLEADMWSIGV 205

Qy       367 CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLS 426
Db       206 ITYILLSGASPFLGDTKQETLANVS-------AVNYDF----------EEEFFRNTSTLA 248

Qy       427 LEMVN---FKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN-EKTKEFLTPLHVAS 482
Db       249 KDFIRRLLVKDPKKRMT-IQDSLQHPWIKPKDTQQALSRKASAVNMEKFKKF------AA 301

Qy       483 EKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR--------------LLLS 528
Db       302 RKKWKQSVRL-------ISLCQRLSRSFLSRS----NMSVARSDDTLDEEDSFVMKAIIH 350

Qy       529 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 588
Db       351 AINDDNVPGLQHL---------LGSLSSYDVNQPNKHGTPPLLIAAGCGNIQMLQLLIKR 401

Qy       589 QS-VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 647
Db       402 GSRIDVQDKGG--SNAIYWASRHGHVDTLKFLNENKCPLDVKDKSGETALHVAARYGHAD 459

Qy       648 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
Db       460 VVQLLCSFGSNPDFQDKEEETPLHCAAWHGYYSVAKALCEVGCNVNIKNREGETP----- 514

Qy       708 DGDTDIQDLLRGDAALLDAAKKG------CLARVKKLSSPDNVNCRDTQGRHSTPLHLAA 761
Db       515 --------------LLTASARGYHDIVECLA-----EHGADLNASDKDGH--IALHLAV 552

Qy       762 GYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFT 821
Db       553 RRCQMEVIKTLLGHGSFVDFQDRHGNTPLHVACKDGSAPIVVALCEASCNLDISNKYGRT 612

Qy       822 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADD---VSALL---------- 868
Db       613 PLHLAANNGILDVVRYLCLMGANVEALTSDGKTAEDLAKAEQHEHVAGLLARLRKDTHRG 672

Qy       869 --------TAAMPP--------------SALPSCYKPQVLNGV---RSPGATADALSSGP 903
Db       673 LFIQQLRPTQNLQPRIKLKLFGHSGSGKSTLVESLKCGLLRSFFRRRRPRLSSTNSTRFP 732

Qy       904 SSP--SSLSAASSLDNLSGSFSELS----SVVSSSG-TEG-------------------- 936
Db       733 PSPLAAKPTVSVSINNLYPGCENVSVRSRSMMFEPGLTKGMLEVFVAPSHHLHCSTDDQS 792

Qy       937 --ASSLEKKEVPGV-DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN- 992
Db       793 TKAIDIQNAYLNGVGDFSVWEFSGNPVYFCCYDYFAANDPTSIHIIVFSLEEPYEIQLNQ 852

Qy       993 ------------------AYGHR------------HKLIKGVERLISGQQGL-------- 1014
Db       853 VIFWLSFLKSLVPVEEPIAFGGKLKNPLRVVLVATHADIMNIPRPAGGEFGYDKDTSLLK 912

Qy      1015 ------------NPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV 1050
Db       913 EIRNRFGNDLHVSNKLFVLDAGASG------SKDIKVLRNHLQEIRSQI 955
```

RESULT 20
ASB2_HUMAN
ID   ASB2_HUMAN   STANDARD;   PRT;   587 AA.
AC   Q96Q27; Q9NSU5; Q9Y567;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 2 (ASB-2).
GN   ASB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND INDUCTION.
RX   MEDLINE=21450655; PubMed=11566180;
RA   Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,
RA   Nakanishi T., Tanaka K.;
RT   "ATRA-regulated Asb-2 gene induced in differentiation of HL-60
RT   leukemia cells.";
RL   FEBS Lett. 505:223-228(2001).
RN   [2]
RP   SEQUENCE FROM N.A., AND INDUCTION.
RX   MEDLINE=21634875; PubMed=11682484;
RA   Guibal F.C., Moog-Lutz C., Smolewski P., Di Gioia Y.,
RA   Darzynkiewicz Z., Lutz P.G., Cayre Y.E.;
RT   "ASB-2 inhibits growth and promotes commitment in myeloid leukemia
RT   cells.";
RL   J. Biol. Chem. 277:218-224(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 31-587 FROM N.A.
RC   TISSUE=Uterus;
RA   Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA   Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 42-587 FROM N.A.
RX   MEDLINE=20564172; PubMed=11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2000).
CC   -!- FUNCTION: Not known. Its expression induces growth inhibition and
CC       chromatin condensation.
CC   -!- INDUCTION: By all-trans retinoic acid (ATRA).
CC   -!- SIMILARITY: Contains 12 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB056723; BAB64532.1; -.
DR   EMBL; AJ251238; CAC17765.1; -.
DR   EMBL; BC032354; AAH32354.1; -.
DR   EMBL; AL137735; CAB70899.1; -.
DR   EMBL; AF159164; AAD45345.1; -.
DR   PIR; T46507; T46507.
DR   Genew; HGNC:16012; ASB2.
DR   MIM; 605759; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 11.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       56     85       ANK 1.
FT   REPEAT       89    119       ANK 2.
FT   REPEAT      123    152       ANK 3.
FT   REPEAT      156    185       ANK 4.
FT   REPEAT      189    218       ANK 5.
FT   REPEAT      222    251       ANK 6.
FT   REPEAT      255    284       ANK 7.
FT   REPEAT      288    317       ANK 8.
FT   REPEAT      320    349       ANK 9.
FT   REPEAT      362    391       ANK 10.
FT   REPEAT      392    421       ANK 11.
FT   REPEAT      428    456       ANK 12.
FT   DOMAIN      533    587       SOCS BOX.
FT   CONFLICT    224    224       MISSING (IN REF. 5).
FT   CONFLICT    460    460       MISSING (IN REF. 5).
SQ   SEQUENCE   587 AA;  65083 MW;  866A2B6010C4C1F9 CRC64;

  Query Match             5.5%;  Score 357;  DB 1;  Length 587;
  Best Local Similarity   23.7%;  Pred. No. 1.4e-13;
  Matches  146;  Conservative   80;  Mismatches  215;  Indels  174;  Gaps   18;

Qy        102 LFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEY--LLQNGANVQ 159
              || :|                || ||  :|:|     |  :  : :|   |   : :
Db         12 LFHSC--------------SAPSRSTAPPESSPARAPMGLFQGVMQKYSSSLFKTSQLA 56

Qy        160 ARDDGGLIPLHNACSFGHAEVVNLLLRHG---ADPNARDNWNYTPLHEAAIKGKIDVCIV 216
               |    ||  |   |  | :  ::: |    |:|| :: |   ||||||  |::  |:
Db         57 PAD-----PLIKAIKDGDEEALKTMIKEGKNLAEPN-KEGW--LPLHEAAYYGQVG-CLK 107

Qy        217 LLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNV 276
               :||    ||  | ||   | : :|  : |     | ||   ::| | :
Db        108 VLQRAYPGTI---DQRT---LQEETAVYLATCRGHLDCLLSLLQAGAEPDI--------- 152

Qy        277 NCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTEL 336
                |:  : |||: |      : |::|:|| || : :      || : |   || ::
Db        153 ----SNKSRETPLYKACERKNAEAVKILVQHNADTNHRCNRGWTALHESVSRNDLEVMQI 208

Qy        337 LVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQL 396
               ||  || | : : :  |||  ||   ::|     |  ||||         | ||
Db        209 LVSGGAKVESKNAYGITPLFVAAQSGQLEALRFLAKYGADINTQASDNASA--------- 259

Qy        397 KERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQI 456
                        | :| :                                    : :::
Db        260 ------------LYEACKN-----------------------------------EHEEV 271

Qy        457 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 516
               | || :||: |:  |: | |||:||:|  :  :|::::   ::   :   | : || ||
Db        272 VEFLLSQGADANKTNKDGLLPLHIASKKGNYRIVQMLLPVTSRTR-IRRSGVSPLHLAAE 330

Qy        517 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 576
                 |:   |||   | |
Db        331 RNHDEVLEALLSARFDVN---------------------------------------- 348

Qy        577 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
                        :     |  | |:|: |:||   | |   | |||||||   ::  : |
Db        349 ---------TPLAPERARLYEDRRSSALYFAVVNNNVYATELLLQHGAD---PNRDVISP 396

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA-------AKGKYEICKLLLQHG 689
               |  |  :|     :||: ||| ::      :   |  |       |     : | |:  |
Db        397 LLVAIRHGCLRTMQLLLDHGANIDA-----YIATHPTAFPATIMFAMKCLSLLKFLMDLG 451

Qy        690 AD--PTKKNRDGNTP 702
               |  |      || |
Db        452 CDGEPCFSCLYGNGP 466


RESULT 21
ANR5_HUMAN
ID   ANR5_HUMAN     STANDARD;      PRT;   776 AA.
AC   Q9NU02; Q9H6Y9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat domain protein 5.
GN   ANKRD5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
```

```
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [2]
RP   SEQUENCE OF 329-776 FROM N.A.
RC   TISSUE=Colon;
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA   Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL109754; CAB87618.2; -.
DR   EMBL; AK025322; BAB15111.1; ALT_INIT.
DR   HSSP; P80144; 2MYO.
DR   Genew; HGNC:15803; ANKRD5.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT      184    213       ANK 2.
FT   REPEAT      217    246       ANK 3.
FT   REPEAT      250    279       ANK 4.
FT   REPEAT      524    553       ANK 5.
FT   REPEAT      557    586       ANK 6.
FT   REPEAT      590    619       ANK 7.
FT   REPEAT      623    652       ANK 8.
SQ   SEQUENCE   776 AA;  86664 MW;  2F71F35AC4D337B6 CRC64;

  Query Match             5.5%;  Score 354.5;  DB 1;  Length 776;
  Best Local Similarity   22.7%;  Pred. No. 2.9e-13;
  Matches  182;  Conservative  117;  Mismatches  303;  Indels  201;  Gaps   30;

Qy       101 ELFEACRNGD---VERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGAN 157
Db        15 KVLQCVRNKDKKQIEKLTKLGYPELINYTEPINGLSA-LHLASVSNDIDMVSFLLDLGAH 73

Qy       158 VQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCI-- 215
Db        74 PDVQDRMGCTPTMRAAELGHELSMEILAKAKADMTIVDNEG---------KGVLFYCILP 124

Qy       216 ---------VLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEK 266
Db       125 TKRHYRCALIALEHGADVNNSTYEGKPIFLRACEDAHDV------KDVCLTFLEKG---- 174

Qy       267 MMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNAC 326
Db       175 ----ANPNAIN--SSTGR--TALMEASREGVVEIVRGILERGGEVNAFDNDRHHAAHFAA 226

Qy       327 SYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKS 386
Db       227 KGGFFDILKLLFAYNGDVGLISINGNTPLHYAAMGGFADCCKYIAQRGCDLKWKNLDHK- 285

Qy       387 AIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAA 446
Db       286 ------TPR---------------AVAKEG-------------GFK-----------AA 299

Qy       447 ASPYPKRKQICELLLRKGA-NINE-----------KTKEFLTPLHVASEK-----AHNDV 489
Db       300 SKEIRRAERIANKLARPGAKNPNPLWALRLHDWSVEREAFLREAFAVLDRGDGSISKNDF 359

Qy       490 VEVVVKHEAKVNALDNLGQTSLHRAAYCG-----------HLQTCRLLLSYG------- 530
Db       360 VMVLEERQDYASSEQLAAIAHLHEKTRGGGVNINEFFKGTRYLNKSFVLGSYGPKKKEKG 419

Qy       531 ---------------CDPNIISLQGFTALQMGN---------ENV-----------QQ 553
Db       420 MGKKGKKGKFVLPLPIC---VIPEYAFPRRQDGGPPYYMIETYKNVTDSSRFNRDHPPEH 476

Qy       554 LLQEGISLGNSEADR---QLLEAAKAGDVETVKKLCTVQS---VNCRDIEGRQSTPLHFA 607
Db       477 PIQDDSVWYIDDSEKVFSNINIITKAGDLASLKK--AFESGIPVDMKD--NYYKTPLMTA 532

Qy       608 AGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKF 667
Db       533 CASGNIDVVKFLLEKGANVNATDNFLWTPLHFACHAGQQDIVELLVESGALIDAASINNS 592

Qy       668 TPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAA 727
Db       593 TPLNRAIESCRLDTVKYLLDIGAKFQLENRKGHSAMDVAK---------AYADYRIIDLI 643

Qy       728 KKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ--HGADVNAQ--- 782
Db       644 KE----KLDNLPKP----AENQKLKGKTPPILKTEGPEIKKEEELLSSIYGVPTTSEGKK 695

Qy       783 -DKGGLIPLHNAASYGH---VDV 801
Db       696 VQKGNVVHLNSLITSGYTKKVDI 718


RESULT 22
V155_FOWPV
ID   V155_FOWPV     STANDARD;      PRT;   542 AA.
AC   Q9J5A7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV115.
GN   FPV115.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44459.1; -.
DR   HSSP; P80144; 2MYO.
```

```
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 2.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       33     62       ANK 1.
FT   REPEAT      157    186       ANK 2.
FT   REPEAT      218    247       ANK 3.
FT   REPEAT      251    281       ANK 4.
FT   REPEAT      285    314       ANK 5.
FT   REPEAT      316    345       ANK 6.
FT   REPEAT      347    375       ANK 7.
FT   REPEAT      378    407       ANK 8.
SQ   SEQUENCE   542 AA;  62056 MW;  CBB71D6E04A06629 CRC64;

  Query Match             5.4%;  Score 351;  DB 1;  Length 542;
  Best Local Similarity   25.9%;  Pred. No. 2.7e-13;
  Matches  137;  Conservative   85;  Mismatches  171;  Indels  136;  Gaps   24;

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQF--------TPLHEAASKNRVEVCSLLLSY 373
              |||     | |:           : |:::::          |||| |    | ::   ||
Db          6 LHNIMYTGTYD---------EIMRAINIYEHRNYSSFRNLPLHYAIYSRRKDIVETLLKS 56

Qy        374 GADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE-ADVTRIKKHLSLEMVNFK 433
              | ||  ::  : : : |   |       ::    ::|     |      |  :||    | |
Db         57 GYDPNSVDITDNNCLQLLSMP-------FDI---TMLPVDEEVQDYAISFYLS---KNMK 103

Qy        434 H-----PQTHETALHCAAASPYP------------KRKQIC--ELLLRKGANINEKTKEF 474
              |     | | |          : ||           | ::|  :|||| ||  | : |:
Db        104 HTSMLIPITKE----ALRGNRYPSEPYFSSMCRKFKDNELCIMDLLLRYGALPNSR-KDG 158

Qy        475 LTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL---LLSYGC 531
              | ||: |:   : ::||::: : || |       :    |    :::| ::   | :|
Db        159 LLPLYHAAAAGNTEMVELLLSYGAKTNLHTRYEDSIFMCAIKSNNVKTAKIISDLYNYKN 218

Qy        532 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 591
              | | |       :|: | ::   |   | :|                          :
Db        219 DINNI----LKTIQLYNADMLLFL---IEIG--------------------------LDI 245

Qy        592 NCRDIEGRQSTPLHFAA-GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAE 650
              | :| :|:  | ||:|    | :  |: ::::|||:: ||: || |||:|| ||  ::::
Db        246 NTKDKKGK--TALHYACNSINCIETVKEIMKYGADINVKDREGLTPLHSACKYGDLKLSK 303

Qy        651 LLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNT----PLDL- 705
              ||:::|| | |       | |: |   |  |: | |::      || :   |     | |
Db        304 LLIEYGADVKVKT--TSTVLNLAVESGNVELVKFLIE-------KNPEFITSDYLSLSLA 354

Qy        706 VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNN 765
              ::  | :|  ||  ||  :|      |:                 ||||||       |
Db        355 IRCKDINIV-LLLLDAG-MDVNSSKCI--------------------STPLHLGVILGN 391

Qy        766 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAAL------LIKY 808
                : : || |||::|| ||:|  ||  |    ::| | |       :|||
Db        392 SNIVKLLLDHGANINAIDKYGETPLEAANKRINIDYAELYKSNRFIIKY 440


RESULT 23
ANR5_MOUSE
ID   ANR5_MOUSE     STANDARD;      PRT;   775 AA.
AC   Q9D2J7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat domain protein 5.
GN   ANKRD5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK019547; BAB31791.1; -.
DR   HSSP; P80144; 2MYO.
DR   MGD; MGI:2441685; Ankrd5.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT      184    213       ANK 2.
FT   REPEAT      217    246       ANK 3.
FT   REPEAT      250    279       ANK 4.
FT   REPEAT      524    553       ANK 5.
FT   REPEAT      557    586       ANK 6.
FT   REPEAT      590    619       ANK 7.
FT   REPEAT      623    652       ANK 8.
SQ   SEQUENCE   775 AA;  86903 MW;  51C11E1BC5151F00 CRC64;

  Query Match             5.4%;  Score 351;  DB 1;  Length 775;
  Best Local Similarity   23.3%;  Pred. No. 4.6e-13;
  Matches  167;  Conservative   92;  Mismatches  271;  Indels  186;  Gaps   25;

Qy        102 LFEACRNGD---VERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANV 158
              : :  || |   :|:: ||  || :| :      |  || |:      |:| :||: ||:
Db         16 VLQCVRNKDKKQIEKLTRLGYPELINFTEPIDGLSA-LHLASISNDTDMVSFLLKLGAHP 74

Qy        159 QARDDGGLIPLHNACSFGHAEVVNLL---------------------------------- 184
                :|  |  |   |  ||   : :|
Db         75 DVQDHMGCTPTMRAAELGHELSMEILAKAKADMTIVDNEGKGVLFYCILPTKRHYRCSLI 134

Qy        185 -LRHGADPNARDNWNY--TPLHEAAIKGKIDV---CIVLLQHGAEPTIRNTD-GRTALDL 237
               | |||| |   | |   |:   | :  ||   |:  |: || |   || ||||
Db        135 ALEHGADVN---NITYEGKPVFLRACEEAHDVKDMCLTFLEKGANPNAINTSTGRTA--- 188

Qy        238 ADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNR 297
                          |:||:| |
Db        189 ----------------LMESSREG---------------------------------V 197
```

```
Qy        298 VKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHE 357
Db        198 LEIVRGILERGGEVNAYDNDRHHAAHFAAKGGFFDILKLLFAYNGDMGLIGMDGNTPLHF 257

Qy        358 AASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREAD 417
Db        258 AAMGGFADCCKYIAQRGCDLKWKNLEHK-------TPRVVAKDG-GFKAAS--KEIRRAE 307

Qy        418 VTRIKKHLSLEMVNFKHPQ------------THETALHCAAASPYPKRKQICELLLRKGA 465
Db        308 RTAAK----LAKTGAKNPNPLWALRLHDWSIEHETSLRNAFK------------FVDRGD 351

Qy        466 NINEKTKEFLTPLHVASEKAHND-VVEVVVKHEAKVNALDNL-----GQTSLHRAAYCGH 519
Db        352 GVVSK-DDFVVALEERQEYATSEQLLSVAQMHEKSRGGGVNINEFFKGTKYLSKSYVLGS 410

Qy        520 LQTCRLLLSYGCDPN-----------IISLQGFTALQMGN---------ENV-------- 551
Db        411 FGPKKKKRGLGKKPRKGKFVLPLPICTIPENAFPRRPDGGPPYYMIETYQNVSDSHRFNR 470

Qy        552 ----QQLLQEGISLGNSEADRQLLEAA---KAGDVETVKK-LCTVQSVNCRDIEGRQSTP 603
Db        471 DHPPEHPIQDDSEWYIDDPSRVFANISFITKAGDLASLKKAIETGIPVDMKD--NTYKTP 528

Qy        604 LHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVAD 663
Db        529 LMIACASGNIDVVKFLIEKGANVNATDNFLWTPLHFACHAGQQDIVELLVKAGASIDATS 588

Qy        664 LWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD-GDTDIQDLLR 718
Db        589 INNSTPLSRAIESCRLDTVKYLLDMGAKFQIENRKGHAAMDIAKAYADYRIIDMIK 644
```

RESULT 24
ANS1_HUMAN
ID   ANS1_HUMAN   STANDARD;   PRT; 1134 AA.
AC   Q92625;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat and SAM domain containing protein 1.
GN   ANKS1 OR KIAA0229.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=97191544; PubMed=9039502;
RA   Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA   Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. VI.
RT   The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT   analysis of cDNA clones from cell line KG-1 and brain.";
RL   DNA Res. 3:321-329(1996).
CC   -!- SIMILARITY: Contains 6 ANK repeats.
CC   -!- SIMILARITY: Contains 1 PID domain.
CC   -!- SIMILARITY: Contains 2 SAM domains.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; D86982; BAA13218.1; ALT_INIT.
DR   HSSP; Q02410; 1AQC.
DR   Genew; HGNC:20961; ANKS1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR006020; PTB_PID.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 7.
DR   Pfam; PF00640; PID; 1.
DR   Pfam; PF00536; SAM; 2.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00462; PTB; 1.
DR   SMART; SM00454; SAM; 2.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS01179; PID; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 2.
KW   ANK repeat; Repeat.
FT   DOMAIN   1   55   GLY-RICH.
FT   REPEAT   79   108   ANK 1.
FT   REPEAT   112   141   ANK 2.
FT   REPEAT   148   177   ANK 3.
FT   REPEAT   181   210   ANK 4.
FT   REPEAT   214   243   ANK 5.
FT   REPEAT   246   275   ANK 6.
FT   DOMAIN   696   762   SAM 1.
FT   DOMAIN   770   837   SAM 2.
FT   DOMAIN   936   1091   PID.
SQ   SEQUENCE   1134 AA; 123081 MW; 21B7171BCE4800BA CRC64;

Query Match 5.3%; Score 341.5; DB 1; Length 1134;
Best Local Similarity 20.8%; Pred. No. 2.9e-12;
Matches 215; Conservative 134; Mismatches 304; Indels 379; Gaps 42;

```
Qy        100 RELFEACRNGDVERVKRLVTPEK---------------------------------- 122
Db          5 QELLEAARTGHLPAVEKLLSGKRLSSGFGGGGGGGSGGGGGGSGGGGGGLGSSSHPLSSL 64

Qy        123 --------VNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACS 174
Db         65 LSMWRGPNVNCVDSTG--YTPLHHAALNGHKDVVEVLLRNDALTNVADSKGCYPLHLAAW 122

Qy        175 FGHAEVVNLLLRHGAD---PNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG 231
Db        123 KGDAQIVRLLIHQGPSHTRVNEQNNDNETALHCAAQYGHTEVVKVLLEELTDPTMRNNKF 182

Qy        232 RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHL 291
Db        183 ETPLDLA---------ALYGRLEVV---------KMLLNAHPNLLSCNT---KKHTPLHL 221

Qy        292 AAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQ 351
Db        222 AARNGHKAVVQVLLDAGMD----------------SNYQ-TEM--------------- 247

Qy        352 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLA---PTPQLKERLAYEFKGHS 408
Db        248 GSALHEAALFGKTDVVQILLAAGTDVNIKDNHGLTALDTVRELPS-QKSQQIAALIEDHM 306

Qy        409 L-LQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 467
Db        307 TGKRSTKEVDKT--------------------------PPPQPP-------LISSMDSI 332

Qy        468 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 527
Db        333 SQKSQGDV-------EKA---VTELIIDFDA------NAEEEGPYEALY----------- 365

Qy        528 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 587
Db        366 ------NAISCHSLDSMASGR-------------SSDQDSTNKEAEAAG---------- 395

Qy        588 VQSVNCRDIEGRQSTPL----------HFAAGYNRVSVVEYLLQ----HGADVHAKDKGG 633
Db        396 VKPAGVRPRE-RPPPPAKPPPDEEEEDHIDKKYFPLTASEVLSMRPRIHGSAAREEDE-- 452

Qy        634 LVPLHNACSYGH-YEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADP 692
Db        453 ----------HPYEL--LLTAETKKVVLVD------------GKTKDHRRSSSSRSQD 486
```

```
Qy        693 TKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGR 752
              : : :||  |           || | : : :  :           |  : |
Db        487 SAEGQDGQVP--------EQFSGLLHGSSPVCEVGQ----------DPFQLLC------ 521

Qy        753 HSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 812
                                                  |   | | :   : ||
Db        522 ------------------------------------TAGQSHPDGSP---QQGACH 538

Qy        813 NATDKWAFTPLHEAAQKGRTQLCAL---------------------LLAHGADPTLKNQ 850
               |: :   | :|     | :|  ||                    | | | |
Db        539 KASMQLEETGVH---APGASQPSALDQSKRVGYLTGLPTTNSRSHPETLTHTASPHPGGA 595

Qy        851 EGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSL- 909
              |          | |   : | ||:: |   | ::         : :| |:  |:  :::
Db        596 E---------EGDRSGARSRA-PPTSKP---KAELKLSRSLSKSDSDLLTCSPTEDATMG 642

Qy        910 SAASSLDNLS-------------GSFSELSSVVSSSGTEGASSLEKKEVPG---VDFSIT 953
              | : || | |                : |:  ::|| |     | |:::: |   :: |:
Db        643 SRSESLSNCSIGKKRLEKSPSFASEWDEIEKIMSSIGEGIDFSQERQKISGSRTLEQSVG 702

Qy        954 QFVRNLGLEH-----LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERL- 1007
              ::: ::||:      |:: |:         ::|   ::|::|||:   || ||::    |
Db        703 EWLESIGLQQYESKLLLNGFDDVHFLGSNVME--EQDLRDIGISDPQHRRKLLQAARSLP 760

Qy       1008 ---ISGQQGLNP 1016
                |  | :|
Db        761 KVKALGYDGNSP 772
```

```
RESULT 25
HMT1_HUMAN
ID   HMT1_HUMAN     STANDARD;      PRT;  1267 AA.
AC   Q9H9B1; Q8TCN7; Q96F53; Q96JF1; Q96KH4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Histone-lysine N-methyltransferase, H3 lysine-9 specific 5
DE   (EC 2.1.1.43) (Histone H3-K9 methyltransferase 5) (H3-K9-HMTase 5)
DE   (Euchromatic histone methyltransferase 1) (Eu-HMTase1) (G9a-like
DE   protein 1) (GLP1).
GN   EUHMTASE1 OR KIAA1876.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), ENZYMATIC ACTIVITY, AND IDENTIFICATION
RP   IN COMPLEX WITH E2F6; TFDP1; MAX; MGA; BAT8; CBX3; RING1; RNF2; MBLR;
RP   L3MBTL2 AND YAF2.
RC   TISSUE=Cervical carcinoma;
RX   MEDLINE=21999559; PubMed=12004135;
RA   Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT   "A complex with chromatin modifiers that occupies E2F- and
RT   Myc-responsive genes in G0 cells.";
RL   Science 296:1132-1136(2002).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Teratocarcinoma;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 351-1267 FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RA   Tsuritani K., Ukai Y., Yajima Y., Amemiya C., Yoshimoto M.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 351-1267 FROM N.A. (ISOFORM 3).
RX   MEDLINE=22158633; PubMed=12168954;
RA   Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT   "Construction of expression-ready cDNA clones for KIAA genes: manual
RT   curation of 330 KIAA cDNA clones.";
RL   DNA Res. 9:99-106(2002).
RN   [5]
RP   SEQUENCE OF 1150-1267 FROM N.A. (ISOFORM 3), AND TISSUE SPECIFICITY.
RC   TISSUE=Brain;
RX   MEDLINE=21245130; PubMed=11347906;
RA   Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XX.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 8:85-95(2001).
RN   [6]
RP   SEQUENCE OF 552-1267 FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SEQUENCE OF 1056-1267 FROM N.A. (ISOFORM 1).
RC   TISSUE=Lymph node;
RA   Ansorge W., Wirkner U., Mewes H.-W., Weil B., Wiemann S.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Histone methyltransferase. Methylates Lys-9 of histone
CC       H3 (in vitro). H3 Lys-9 methylation represents a specific tag for
CC       epigenetic transcriptional repression by recruiting HP1 proteins
CC       to methylated histones. Probably targeted to histone H3 by
CC       different DNA-binding proteins like E2F6, MGA, MAX and/or DP1.
CC       During G0 phase, it probably contributes to silencing of MYC- and
CC       E2F-responsive genes, suggesting a role in G0/G1 transition in
CC       cell cycle.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC       S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC   -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC       E2F6, MGA, MAX, TFDP1, CBX3, BAT8, EUHMTASE1, RING1, RNF2, MBLR,
CC       L3MBTL2 and YAF2.
CC   -!- SUBCELLULAR LOCATION: Nuclear; associates with euchromatic
CC       regions.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Experimental confirmation may be lacking for some
CC         isoforms;
CC       Name=1;
CC         IsoId=Q9H9B1-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9H9B1-2; Sequence=VSP_002222, VSP_002223;
CC       Name=3;
CC         IsoId=Q9H9B1-3; Sequence=VSP_002224, VSP_002225;
CC   -!- TISSUE SPECIFICITY: Widely expressed.
CC   -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC       family.
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   -!- SIMILARITY: Contains 1 pre-SET domain.
CC   -!- SIMILARITY: Contains 1 SET domain.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AY083210; AAM09024.1; -.
DR   EMBL; AK022941; BAB14321.1; -.
DR   EMBL; AB028932; BAB56104.1; -.
DR   EMBL; AB058779; BAB47505.2; -.
DR   EMBL; BC011608; AAH11608.2; -.
DR   EMBL; AL713772; CAD28534.1; ALT_INIT.
DR   MIM; 607001; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR007728; Pre-SET.
DR   InterPro; IPR001214; SET.
DR   InterPro; IPR003606; Zn2-binding.
DR   Pfam; PF00023; ank; 7.
DR   Pfam; PF05033; Pre-SET; 1.
DR   Pfam; PF00856; SET; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 7.
DR   SMART; SM00468; PreSET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50867; PRE_SET; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW   ANK repeat; Repeat; Alternative splicing.
FT   REPEAT      706    735       ANK 1.
FT   REPEAT      741    770       ANK 2.
FT   REPEAT      774    803       ANK 3.
FT   REPEAT      807    837       ANK 4.
FT   REPEAT      841    870       ANK 5.
FT   REPEAT      874    903       ANK 6.
FT   REPEAT      907    936       ANK 7.
FT   REPEAT      940    973       ANK 8.
FT   DOMAIN     1029   1092       PRE-SET.
FT   DOMAIN     1094   1216       SET.
FT   DOMAIN      375    378       POLY-GLU.
FT   DOMAIN      411    418       POLY-ARG.
FT   DOMAIN     1261   1264       POLY-ALA.
FT   VARSPLIC    805    825       DAEGSTCLHLAAKKGHYEVVQ -> IQKTSKVYTESQETQR
FT                                SQTIL (in isoform 2).
FT                                /FTId=VSP_002222.
FT   VARSPLIC    826   1267       Missing (in isoform 2).
FT                                /FTId=VSP_002223.
FT   VARSPLIC   1150   1153       DGEV -> ISSA (in isoform 3).
FT                                /FTId=VSP_002224.
FT   VARSPLIC   1154   1267       Missing (in isoform 3).
FT                                /FTId=VSP_002225.
FT   CONFLICT    530    530       G -> E (IN REF. 3 AND 4).
SQ   SEQUENCE   1267 AA;  138181 MW;  78FD418FD404F475 CRC64;

  Query Match             5.1%;  Score 330;  DB 1;  Length 1267;
  Best Local Similarity   22.6%;  Pred. No. 1.6e-11;
  Matches  183;  Conservative  102;  Mismatches  319;  Indels  204;  Gaps   27;

Qy        220 HGAE-----PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE--------- 265
              ||||     ||  :   : ::  || : |  : || :::: ||  :| ||
Db        331 HGAEQAAAFPTEDSRTSKESMSEADRAQK--MDGESEEEQ--ESVDTGEEEEGGDESDLS 386

Qy        266 --------------------------------KMMALLTPLNVNCHASDGRKSTPLHLA 292
                                             |   |   :   |   :: | :
Db        387 SESSIKKKFLKRKGKTDSPWIKPARKRRRRSRKKPSGALGSESYKSSAGSAEQTAPGD-S 445

Qy        293 AGYNRVKIVQLLLQ------------HGADVHAKDKGDLVPLHNACSY-----GHYEVT 334
              ||  | :  | |:             :| ||   |    |||  ||         |:|
Db        446 TGYMEVSLDSLDLRVKGILSSQAEGLANGPDVLETDGLQEVPL---CSCRMETPKSREIT 502

Qy        335 ELLVKHGACVNAMDLWQFTPLHE-AASKNRVEVCSLLLSYGADPTLLNC--HNKSAIDLA 391
              |     ::|          ||     | |    |:       | |: |   |    :
Db        503 TLANNQCMATESVD-------HELGRCTNSVVKYGLMRPSNKAPLLVLCEDHRGRMVKHQ 555

Qy        392 PTPQLKERLAYEFKGHSLLQAAREADVT-RIKKHLSLEMVNFKHPQTHETALHCAAASPY 450
              |       |     : ::   |: :: | |  :  : | :              ||   |
Db        556 CCP----GCGYFCTAGNFMECQPESSISHRFHKDCASRVNNASY------CPHCGEES-- 603

Qy        451 PKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTS 510
                                 | ||      |  ||       |  : | :||:     |:
Db        604 ------------------SKAKE------VTIAKADTTSTVTPVPGQEKGSALEGRADTT 639

Qy        511 LHRAA-----YCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSE 565
                ||           ||      :  | ||   :  |     :      :: |:  :   :||
Db        640 TGSAAGPPLSEDDKLQGAASHVPEGFDPTGPAGLGRPTPGLSQGPGKETLESALIALDSE 699

Qy        566 A-------DRQLLEAAKAGDVETVKKLCTVQSV--NCRDIEGRQSTPLHFAAGYNRVSVV 616
                    :||  :|: |::: |  |  |  :  | :      : :||| ||     | :
Db        700 KPKKLRFHPKQLYFSARQGELQKV-LLMLVDGIDPNFKMEHQNKRSPLHAAAEAGHVDIC 758

Qy        617 EYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAK 676
                |:| ||::     :    ||  |    | |  : |:| ||:|:  |     | || ||| |
Db        759 HMLVQAGANIDTCSEDQRTPLMEAAENNHLEAVKYLIKAGALVDPKDAEGSTCLHLAAKK 818

Qy        677 GKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVK 736
              | ||: : || :|
Db        819 GHYEVVQYLLSNG---------------------------------------------- 831

Qy        737 KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 796
                  :|||:|  |   ||:  |  | :::: : ||  |:|:| :|    | || ||
Db        832 ----QMDVNCQDDGG--WTPMIWATEYKHVDLVKLLLSKGSDINIRDNEENICLHWAAFS 885

Qy        797 GHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856
              | ||:| :|:     ::| :    :||| ||:: |      | |:  :| ||||:||:|||
Db        886 GCVDIAEILLAAKCDLHAVNIHGDSPLHIAARENRYDCVVLFLSRDSDVTLKNKEGETPL 945

Qy        857 DLV-------SADDVSALLTAAMP--PSALPSCYKPQVLNGV-RSPGATADALSSGPSSP 906
                      ||  :|  |  : |  || :       :  |  | |    :|: | |  |
Db        946 QCASLNSQVWSALQMSKALQDSAPDRPSPVERIVSRDIARGYERIPIPCVNAVDSEP-CP 1004

Qy        907 S-----SLSAASSLDNLSGSFSELSSVV 929
              |     | :  :|  |:  : : |   |
Db       1005 SNYKYVSQNCVTSPMNIDRNITHLQYCV 1032


RESULT 26
V228_FOWPV
ID   V228_FOWPV     STANDARD;      PRT;   525 AA.
AC   Q9J507;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV228.
GN   FPV228.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44572.1; -.
DR   HSSP; P25963; 1IKN.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 9.
DR   SMART; SM00248; ANK; 9.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       39     71       ANK 1.
FT   REPEAT       72    122       ANK 2.
FT   REPEAT      123    152       ANK 3.
FT   REPEAT      156    185       ANK 4.
FT   REPEAT      190    226       ANK 5.
FT   REPEAT      227    254       ANK 6.
FT   REPEAT      258    287       ANK 7.
FT   REPEAT      291    320       ANK 8.
FT   REPEAT      324    353       ANK 9.
SQ   SEQUENCE   525 AA;  60155 MW;  E8AF61D9B33BCFB0 CRC64;

  Query Match             5.1%;  Score 327.5;  DB 1;  Length 525;
  Best Local Similarity   24.4%;  Pred. No. 6.2e-12;
  Matches  122;  Conservative   70;  Mismatches  176;  Indels  133;  Gaps   14;

Qy        322 LHNACSYG-HYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              |::| |    : |  :| |      :: ::      || ||      ::   || ||
Db         10 LYHAVSINDRFSVVNILEKKNIPIDYINFHPDNPLLEAVKLTNTDMIKTLLDYGI----- 64

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR-IKKHLSLEMVNFKHPQTHE 439
              | |                                   || |  : :| ::   :   |
Db         65 -CIN---------------------------------TRDILGNTALHLIAMDYYVPHN 89

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              |      |   |  | ||| ||||       ||||:|:|  :  :::::: : ||
Db         90 DIKHGHHNDYVFKMVPIINLFLRKKANINACNNLNQTPLHLAAESNNTTLLKILLYNNAK 149

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              || ||  | | || |    :::: :|||||  | ||   : ||           | | :
Db        150 VNILDIYGNTCLHYAVRGRNIESIKLLLSYNVDVNI---RNFTYW-------YSALHEAV 199

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVN--CRDIEGRQSTPLHFAAGYNRVSVVE 617
              :|:|:  | ::             ||   :|| ||      :||: :|   | :  ::
Db        200 QIGDSKISRCIVSLL----------LCNKANVNTRCR----LNTTPIFYA--INCIDTLK 243

Qy        618 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 677
              ||::|||::|       :| |      |:: : |: :|| ||:      |||: |
Db        244 LLLENGADINATSDNDNAVIHLATENRRYDIIKTLLDYGADVNMIGYRGKTPLYYATENY 303

Qy        678 KYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKK 737
              |   |||| ||::|   :   |||| :
Db        304 SYRNMKLLLDHGSNPNIADHIMNTPLFI-------------------------------- 331

Qy        738 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPL-HNAASY 796
                ::  |                  :|  : ||  |||:|  :  |  |: : |  :
Db        332 -----SIKC----------------TCIENTKMLLDSGADINHVNDNGETPISYLAPNL 369

Qy        797 GHVDVAALLI--------KYN 809
                || |:|      |||
Db        370 IPTPVAILVISHIVLLKTKYN 390


RESULT 27
ASB3_MOUSE
ID   ASB3_MOUSE     STANDARD;      PRT;   525 AA.
AC   Q9WV72;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 3 (ASB-3).
GN   ASB3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=20564172; PubMed=11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2000).
CC   -!- TISSUE SPECIFICITY: Widely expressed; highest expression in testis
CC       and spleen.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF155354; AAD38810.1; -.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:1929749; Asb3.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       9     38       ANK 1.
FT   REPEAT      42     71       ANK 2.
FT   REPEAT      78    107       ANK 3.
FT   REPEAT     111    140       ANK 4.
FT   REPEAT     145    174       ANK 5.
FT   REPEAT     178    207       ANK 6.
FT   REPEAT     211    240       ANK 7.
FT   REPEAT     246    275       ANK 8.
FT   REPEAT     279    308       ANK 9.
FT   REPEAT     315    346       ANK 10.
FT   DOMAIN     442    505       SOCS BOX.
SQ   SEQUENCE   525 AA;  58354 MW;  FE67680AD3958FC1 CRC64;

  Query Match             5.0%;  Score 322;  DB 1;  Length 525;
  Best Local Similarity   29.7%;  Pred. No. 1.3e-11;
  Matches  118;  Conservative   48;  Mismatches  127;  Indels  104;  Gaps   15;

Qy        573 AAKAGDVETVKKLCTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGAD---VHA 628
              ||: |:|: ::||    :||:  |  |     |:| || :| |  :: |:   :    :|
Db         17 AAREGNVKILRKLLKKGRSVDVADNRGWM--PIHEAAYHNAVECLQMLIHTDSSENYIKA 74

Qy        629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 688
              |   |   || | | ||::: ::|:: ||   |  : |||  |   |: :: ||||||
Db         75 KTFEGFCALHLAVSQGHWKITQILLEAGADPNETTLEETTPLFLAVESGRIDVLKLLLQH 134

Qy        689 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 748
              ||                                                  |||
Db        135 GA-------------------------------------------------NVN--- 139

Qy        749 TQGRHS----TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAAL 804
                | ||      || |:  | |  : ||: |||   ||  |: ||  || || ::  ::
Db        140 --GSHSMSGWNSLHQASFQGNAETIKLLLKQGADRECQDDFGITPLFVAAQYGKLESMSI 197

Qy        805 LIKYNACVN--ATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTL----------KN 849
              ||   | ||  | ||   |||  |||:| |:   |||: |||| |
Db        198 LISSGANVNCQALDK--ATPLFIAAQEGHTKCVELLLSSGADPDLYCNEDNWQLPIHAAA 255
```

```
Qy        850 QEGQT-PLDLV----------SADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADA 898
Db        256 QMGHTETLDLLIPRTNRACDTEPDKVSPVYSAVF--GGREECLEMLLQNGY-SPDAQM-C 311

Qy        899 LSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 935
Db        312 LVFGFSSPLCMAFQKDCD-----FSGIVNILLKYGAQ 343
```

RESULT 28
V245_FOWPV
ID V245_FOWPV STANDARD; PRT; 436 AA.
AC Q9J4Z5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ankyrin-repeat protein FPV245.
GN FPV245.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- SIMILARITY: Contains 10 ANK repeats.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF198100; AAF44589.1; -.
DR HSSP; P80144; 2MYO.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 9.
DR SMART; SM00248; ANK; 8.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 1 30 ANK 1.
FT REPEAT 34 63 ANK 2.
FT REPEAT 67 96 ANK 3.
FT REPEAT 98 119 ANK 4.
FT REPEAT 123 152 ANK 5.
FT REPEAT 156 185 ANK 6.
FT REPEAT 189 218 ANK 7.
FT REPEAT 222 252 ANK 8.
FT REPEAT 253 283 ANK 9.
FT REPEAT 287 317 ANK 10.
SQ SEQUENCE 436 AA; 49509 MW; 7F74EFE846B876C0 CRC64;

```
  Query Match             5.0%;  Score 321;  DB 1;  Length 436;
  Best Local Similarity   27.0%;  Pred. No. 1.1e-11;
  Matches  100;  Conservative   60;  Mismatches  148;  Indels   62;  Gaps    7;

Qy        458 ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 517
Db         19 KLIKNKGNCINISVEETTTPLIDAIRTGNAKIVELFIKHGAQVNHVNTKIPNPLLTAIKI 78

Qy        518 GHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAG 577
Db         79 GSHDIVKLLLLINGVDTSILPVP-----CINKEMIKTILDSGVKVNTKNA---------- 122

Qy        578 DVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPL 637
Db        123 --------------------KSKTFLHYAIKNNDLEVIKMLFEYGADVNIKDDNGCYPI 161

Qy        638 HNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNR 697
Db        162 HIATRSNSYEIIKLLLEKGAYANVKDNYGNSPLHNAAKYGDYACIKLVLDHTNNISNKCN 221

Qy        698 DGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPL 757
Db        222 NGVTPLH---------------NAILYNR------SAVELLINNRSINDTDVDG--YTPL 258

Qy        758 HLAAGYN-NLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD-VAALLIKYNACVNAT 815
Db        259 HYALQPPCSIDIIDILLYNNADISIKDNNGRNPIDTAFKYINRDSVIKELLANAVLINEV 318

Qy        816 DKWAFTPLHE 825
Db        319 GKLKDTTILE 328
```

RESULT 29
V246_FOWPV
ID V246_FOWPV STANDARD; PRT; 592 AA.
AC Q9J4Z4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative ankyrin-repeat protein FPV246.
GN FPV246.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- SIMILARITY: Contains 10 ANK repeats.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF198100; AAF44590.1; -.
DR HSSP; P80144; 2MYO.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 8.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 8.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 63 93 ANK 1.
FT REPEAT 162 191 ANK 2.
FT REPEAT 195 224 ANK 3.
FT REPEAT 227 255 ANK 4.
FT REPEAT 256 284 ANK 5.
FT REPEAT 288 317 ANK 6.
FT REPEAT 321 348 ANK 7.
FT REPEAT 352 381 ANK 8.
FT REPEAT 386 413 ANK 9.
FT REPEAT 417 449 ANK 10.
SQ SEQUENCE 592 AA; 66850 MW; 4529996B135848DC CRC64;

```
  Query Match             4.9%;  Score 318;  DB 1;  Length 592;
  Best Local Similarity   21.5%;  Pred. No. 2.7e-11;
  Matches  131;  Conservative   89;  Mismatches  213;  Indels  176;  Gaps   20;

Qy        145 KDVVEYLLQNGANVQARDDGGLI--------PLHNACSFGHAEVVNLLLRHG--ADPNAR 194
```

```
Db         36 RDRIEEIIELNRIIELDDDDVIYDQIYFPYNPLHQVIEARRANLVEIILDQGKYTVNSVN 95

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254

Db         96 DDFIFYPLH-------VLTCV------------------------PETSDVLT--YL--- 119

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314

Db        120 --------NEIEEFTLVKDLEL--------RACKMNLSKSISIAIIKQVL----KGIREF 159

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374

Db        160 TDNDLRKLDNQIREDELKIAQLLISRGAKLDIKNEYGYTPLRNTVINGNFALTKLLLDNG 219

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434

Db        220 ADAS-IKCNGMTIFEISTLSQ--------------------NVDMIKEIIK------RC 251

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494

Db        252 GYNHDSKILCRVAS--KGYINVIHFLLDIGFNVNSFDSFGETPLHAATRSGSVETVNALI 309

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554

Db        310 SYGSTVDIKDNIGSTPLMHA--CGNKDISKLLIEKGADPNISNIHGY------------- 354

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614

Db        355 ------------------------------------------------TPLHNAAAYGSVD 367

Qy        615 VVEYLLQHGADVHAKDK-GGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEA 673

Db        368 VVNLLLSYGVSIDAKDKIIGSTPLDHGVH--HPEIVKLLLEKGANPNTINLYGYTPLKNA 425

Qy        674 AAKGK----YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKK 729

Db        426 ILKSRLSAEYIIPYVVLQDFSILDGRNMPGS--------GIN--MELIRND-KLLQNIKT 474

Qy        730 GCLARVKKL 738

Db        475 SCEIELKRM 483


RESULT 30
FEM1_CAEEL
ID   FEM1_CAEEL     STANDARD;      PRT;   656 AA.
AC   P17221;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sex-determining protein fem-1.
GN   FEM-1 OR ISX-1 OR F35D6.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=90199879; PubMed=2317869;
RA   Spence A.M., Coulson A., Hodgkin J.;
RT   "The product of fem-1, a nematode sex-determining gene, contains a
RT   found in cell cycle control proteins and receptors for
RT   cell-cell interactions.";
RL   Cell 60:981-990(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Sex-determining protein. It is essential for the
CC       adoption of the male sexual fate in all tissues in C.elegans.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03172; AAA28055.1; -.
DR   EMBL; U52001; AAA96093.1; -.
DR   PIR; A34793; A34793.
DR   HSSP; Q00420; 1AWC.
DR   WormPep; F35D6.1; CE07175.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 2.
KW   Developmental protein; Phosphorylation; ANK repeat; Repeat.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT       88    118       ANK 2.
FT   REPEAT      122    151       ANK 3.
FT   REPEAT      155    184       ANK 4.
FT   REPEAT      188    217       ANK 5.
FT   REPEAT      220    250       ANK 6.
FT   REPEAT      557    588       ANK 7.
SQ   SEQUENCE   656 AA;  74286 MW;  4F5FC3FA3EA950D3 CRC64;

  Query Match             4.9%;  Score 315.5;  DB 1;  Length 656;
  Best Local Similarity   23.0%;  Pred. No. 4.4e-11;
  Matches  184;  Conservative   95;  Mismatches  269;  Indels  251;  Gaps   31;

Qy        102 LFEACRNGDVERVKRLVTPEKVNSR------------DTAGRKSTPLHFAAGFGRKDVVE 149

Db         12 IYNAAAVGNLQRIKVFT----INSRNDRQWIIDCFNSDQDGR--YPLVIAARNGHANVVE 65

Qy        150 YLLQNGANVQAR--------DDGGLIPLHNACSFGHAEVVNLLL-RHGADPNARDNWNYT 200

Db         66 YLLEIGADPSVRGVVEFDNENIQGTPPLWAASAAGHIEIVKLLIEKANADVNQATNTRST 125

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260

Db        126 PLRGACYDGHLDIVKYLLEKGADPHIPNRHGHTCLMIA----------SYR--------- 166

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320

Db        167 ---------------------------------NKVGIVEELLKTGIDVNKKTERGNT 191

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380

Db        192 ALHDAAESGNVEVVKILLKHGS-VLMKDIQGVDPLMGAALSGFLDVLNVL----ADQMPS 246

Qy        381 NCHNKSAIDLAPTPQLKERL----AYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQ 436

Db        247 GIHKRDALKLLGSTLLDKKMDAMSAMNCWKQS-MEVPLHADDLRLVR----EMETFFEPQ 301

Qy        437 THETALHCAAASPYPKRKQ-ICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVK 495

Db        302 ---------EVYEYQREAQNIAQVELLDG-NIEAQRMQALVIRERILGGAHTDVHYYL-- 349

Qy        496 HEAKVNALDNLGQTSLHRAAYC--GHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQ 553

Db        350 -------------RFRGAVYCDMGQMNRCYDLWKHALE---LQQKHFAPLYYGTITTLQ 392

Qy        554 LLQEGISL------GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFA 607

Db        393 SFHETFSMSLNDFVNNHHANRNLRV---------------------------------- 417

Qy        608 AGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKF 667
```

```
                 | | |:|:           |  : |  | :                        |
Db        418 ----RSSWVKYVF----------NGVCLELERAAA---------------------WTG 441

Qy        668 TPLHEAAAKGKYEIC-KLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDA 726
              || |       | | |   ||  : ::  :     : ||     :: : |   :| :|
Db        442 APLLEDT-----ECCGKERCQHATEESEYKKLVYVAVHLV-----NVLERLSLPSAHIDD 491

Qy        727 A--KKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 784
              :  :|    | |::|     : | :         ||:   :: ||
Db        492 SDEEKEPKADVRRLM----IVCHE--------LHIPLLHHTLEER--------------- 524

Qy        785 GGLIPLHNAASYGHVDVAAL--LIKYNACVNATDKWAFTPLH--EAAQKGRTQLCALLLA 840
                ||  |:|  |     | |  |:: :  ||||||    ||:|     |:: |  |   ||
Db        525 ---IPDSNSAELGLPKAAVLEQLLELDLDVNATDKNNDTPMHILLRAREFRKSLVRALLV 581

Qy        841 HGADPTLKNQEGQTPLDLV 859
                |     :|: |   |:::
Db        582 RGTWLFARNRHGDVVLNVM 600
```

RESULT 31
PPGB_MOUSE
ID PPGB_MOUSE STANDARD; PRT; 568 AA.
AC Q8VHQ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
DE inhibited membrane-associated protein) (CAAX box protein TIMAP).
GN PPP1R16B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Cerebellum;
RX MEDLINE=22525552; PubMed=12638124;
RA Magdaleno S., Northcutt G.M., Curran T., Kurschner C.;
RT "mPPP1R16B is a novel mouse protein phosphatase 1 targeting subunit
RT whose mRNA is located in cell bodies and dendrites of neurons in four
RT distinct regions of the brain.";
RL Gene Expr. Patterns 1:143-149(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be a downstream target for TGF-beta1 signaling
CC cascade in endothelial cells (By similarity).
CC -!- SUBUNIT: Binds PP1 (Probable).
CC -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (By
CC similarity).
CC -!- MISCELLANEOUS: Repressed by TGF-beta1 which requires denovo
CC protein synthesis and histone deacetylase (By similarity).
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF423761; AAL62093.1; -.
DR EMBL; BC054764; AAH54764.1; -.
DR EMBL; BC057542; AAH57542.1; -.
DR MGD; MGI:2151841; Ppp1r16b.
DR GO; GO:0005886; C:plasma membrane; ISS.
DR GO; GO:0019903; F:protein phosphatase binding; ISS.
DR GO; GO:0007165; P:signal transduction; ISS.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 5.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT DOMAIN 15 55 COILED COIL (POTENTIAL).
FT REPEAT 100 129 ANK 1.
FT REPEAT 133 162 ANK 2.
FT REPEAT 228 257 ANK 3.
FT REPEAT 261 290 ANK 4.
FT REPEAT 531 560 ANK 5.
FT LIPID 565 565 S-farnesyl cysteine (Potential).
SQ SEQUENCE 568 AA; 63571 MW; 5D8446C81DEBA3D1 CRC64;

Query Match 4.8%; Score 312.5; DB 1; Length 568;
Best Local Similarity 29.4%; Pred. No. 5.4e-11;
Matches 96; Conservative 56; Mismatches 132; Indels 43; Gaps 12;

```
Qy        126 RDTAGRKSTPLHFAAGFG------RKDV--VEYLLQNGANVQARDDGGLIPLHNACSFGH 177
              | | ||:   : | |           | |   | | |:|  :    :: ||  ||  |
Db         55 RSTGGRRK-KVSFEASVALLEASLRNDAEEVRYFLKNKVSPDLCNEDGLTALHQCCIDNF 113

Qy        178 AEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDL 237
              |:| ||| |||: ||:|| :|||| ||  | |::  :|:|:||:    |:||    ||
Db        114 EEIVKLLLSHGANVNAKDNELWTPLHAAATCGHINLVKILVQYGADLLAVNSDGNMPYDL 173

Qy        238 A--DPSAKAVLT-GEYK--KDELLESARSGNEEKMMALLTPLNVNCHAS--------DGR 284
                :|:   : |   |:   | :   |: |:||::    :::|  :          ||:
Db        174 CEDEPTLDVIETCMAYQGITQEKINEMRAAPEQKMIS-----DIHCMIAAGQDLDWIDGQ 228

Qy        285 KSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACV 344
              :| ||:|      ::  :||| ||  |  ||      ||| |  :|  : |||| ||| :
Db        229 GATLLHIAGANGYLRAAELLLDHGVRVDVKDWDGWEPLHAAAFWGQMPMAELLVSHGASL 288

Qy        345 NAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN-CHNKSAID-----------LAP 392
              :|       |: :   :  :|  | | :  |  : :   :||::           :
Db        289 SARTSMDEMPI-DLCEEEEFKVLLLELKHKHDVIMKSQLRHKSSLSRRTSSAGSRGKVVR 347

Qy        393 TPQLKER---LAYEFKGHSLLQAAREA 416
                | :|      |::| ::|   | |
Db        348 RASLSDRTNLYRKEYEGEAILWQQRSA 374
```

RESULT 32
PPGA_HUMAN
ID PPGA_HUMAN STANDARD; PRT; 528 AA.
AC Q96I34;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein phosphatase 1 regulatory inhibitor subunit 16A.
GN PPP1R16A.
OS Homo sapiens (Human).

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RT   "The nucleotide sequence of a long cDNA clone isolated from human
RT   spleen.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung, Placenta, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Inhibits protein phosphatase 1 activity toward
CC       phosphorylase, myosin light chain and myosin substrates (By
CC       similarity).
CC   -!- SUBUNIT: Binds PP1.
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (By
CC       similarity).
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   -!- CAUTION: Ref.1 sequence may be incorrect due to a cloning
CC       artifact.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK090471; BAC03452.1; ALT_SEQ.
DR   EMBL; BC007854; AAH07854.1; -.
DR   EMBL; BC049841; AAH49841.1; -.
DR   EMBL; BC053506; AAH53506.1; -.
DR   Genew; HGNC:14941; PPP1R16A.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 4.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   DOMAIN       18     45       COILED COIL (POTENTIAL).
FT   REPEAT       70     99       ANK 1.
FT   REPEAT      103    132       ANK 2.
FT   REPEAT      136    165       ANK 3.
FT   REPEAT      231    260       ANK 4.
FT   REPEAT      264    293       ANK 5.
FT   LIPID       525    525       S-farnesyl cysteine (Potential).
FT   CONFLICT     34     71       MISSING (IN REF. 1).
FT   CONFLICT    249    249       A -> P (IN REF. 1).
FT   CONFLICT    272    272       A -> P (IN REF. 1).
SQ   SEQUENCE   528 AA;  57811 MW;  A5A2D5E1A16A438C CRC64;

  Query Match             4.8%;  Score 310;  DB 1;  Length 528;
  Best Local Similarity   26.2%;  Pred. No. 6.8e-11;
  Matches  116;  Conservative    55;  Mismatches  164;  Indels  108;  Gaps   10;

Qy        553 QLLQEGISLGNSEADRQLLEA------------------AKAGDVETVKKLCTVQSVNC 593
              :|| |   :|    :|  |                      | |  |  :|    | :
Db          6 ELLAEMPMVGRMSTQERLKHAQKRRAQQVKMWAQAEKEAQGKKGPGERPRKEAASQGL-L 64

Qy        594 RDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLV 653
              : :    |  |  ||  | :  |  |  |       :: ||  ||  |      |: : |:
Db         65 KQVLFPPSVVLLEAAARNDLEEVRQFLGSGVSPDLANEDGLTALHQCCIDDFREMVQQLL 124

Qy        654 KHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT-D 712
              : || :|  |   :|||| ||  |   : :||:  ||:    | ||| | ||  |  | |
Db        125 EAGANINACDSECWTPLHAAATCGHLHLVELLIASGANLLAVNTDGNMPYDLCDDEQTLD 184

Qy        713 IQDLLRGDAAL----LDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 768
                :    |  :    ::||:      | :|   |::  |
Db        185 CLETAMADRGITQDSIEAAR-----AVPELRMLDDIRSR---------------------- 218

Qy        769 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQ 828
                || |||::|     |   || ||: |  : ||||::: | ::| |:  : ||| ||
Db        219 ----LQAGADLHAPLDHGATLLHVAAANGFSEAAALLLEHRASLSAKDQDGWEPLHAAAY 274

Qy        829 KGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG 888
              |:  |  ||:|||||    |:   :||||:   ::| | |        |   :|
Db        275 WGQVPLVELLVAHGADLNAKSLMDETPLDVCGDEEVRAKL--------LELKHK------ 320

Qy        889 VRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGV 948
                |||    |   ||                |    ||:|: |
Db        321 -------HDALLRAQSRQRSL--------------LRRRTSSAGSRG------------ 346

Qy        949 DFSITQFVRNLGLEHLMDIFERE 971
                : || : |    |:: ::
Db        347 -----KVVRRVSLTQRTDLYRKQ 364


RESULT 33
PPGA_MOUSE
ID   PPGA_MOUSE     STANDARD;      PRT;   524 AA.
AC   Q923M0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16A (Myosin
DE   phosphatase targeting subunit 3) (Mypt3).
GN   PPP1R16A OR MYPT3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND FARNESYLATION.
RX   MEDLINE=21234869; PubMed=11336659;
RA   Skinner J.A., Saltiel A.R.;
RT   "Cloning and identification of MYPT3: a prenylatable myosin targetting
RT   subunit of protein phosphatase 1.";
RL   Biochem. J. 356:257-267(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
```

```
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Inhibits protein phosphatase 1 activity toward
CC       phosphorylase, myosin light chain and myosin substrates.
CC   -!- SUBUNIT: Binds PP1.
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (By
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis, followed by liver,
CC       heart, brain and lung.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY010723; AAG40949.1; -.
DR   EMBL; BC057450; AAH57450.1; -.
DR   MGD; MGI:1920312; Ppp1r16a.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   REPEAT       70    99       ANK 1.
FT   REPEAT      103   132       ANK 2.
FT   REPEAT      136   165       ANK 3.
FT   REPEAT      231   260       ANK 4.
FT   REPEAT      264   293       ANK 5.
FT   DOMAIN      448   469       COILED COIL (POTENTIAL).
FT   LIPID       521   521       S-farnesyl cysteine (Probable).
SQ   SEQUENCE   524 AA;  57529 MW;  A0F682A794475429 CRC64;

  Query Match             4.8%;  Score 309;  DB 1;  Length 524;
  Best Local Similarity   29.4%;  Pred. No. 7.7e-11;
  Matches  109;  Conservative   42;  Mismatches  142;  Indels   78;  Gaps    9;

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db         72 SVALLEAAARNDLEEVRQFLTSGVSPNLANEDGLTALHQCCIDDFQEMAQQLLDAGADVN 131

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
Db        132 ARDSECWTPLHAAATCGHLHLVELLISRGADLLAVNSDGNMPYDLCEDAQT--LDCLETA 189

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
Db        190 MANQGITQEGI---EEARAVPELCMLNDLQNR------LQAGANLSDP----LDHGATL- 235

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
Db        236 ---------LHIAAANGFSEVATLLLEQGASLSAKDHDGWEPLHAAAYWGQVHLVELLVA 286

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
Db        287 HGADLNGKSLVDETPLDVCGDEEVRAKL--------LELKHK------------QDALL 325

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
```

```
                 ||              |    ||:|: |                : || :
Db        326 RAQGRQRSL--------------LRRRTSSAGSRG----------------KVVRRVS 353

Qy        961 LEHLMDIFERE 971
              | |  ::: :|
Db        354 LTHRTNLYRKE 364


RESULT 34
ASB3_HUMAN
ID   ASB3_HUMAN     STANDARD;      PRT;   518 AA.
AC   Q9Y575; Q9NVZ2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 3 (ASB-3).
GN   ASB3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20564172; PubMed=11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 11 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF156778; AAD41895.1; -.
DR   EMBL; AK000985; BAA91455.1; -.
DR   EMBL; AK001283; BAA91599.1; -.
DR   EMBL; BC006488; AAH06488.1; -.
DR   EMBL; BC009569; AAH09569.1; -.
DR   Genew; HGNC:16013; ASB3.
DR   MIM; 605760; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT        9     38       ANK 1.
FT   REPEAT       42     71       ANK 2.
FT   REPEAT       78    107       ANK 3.
FT   REPEAT      111    140       ANK 4.
FT   REPEAT      145    174       ANK 5.
FT   REPEAT      178    207       ANK 6.
FT   REPEAT      211    240       ANK 7.
FT   REPEAT      246    275       ANK 8.
FT   REPEAT      279    308       ANK 9.
FT   REPEAT      315    346       ANK 10.
FT   REPEAT      348    373       ANK 11.
FT   DOMAIN      441    504       SOCS BOX.
FT   CONFLICT    118    118       L -> S (IN REF. 2; BAA91599).
SQ   SEQUENCE   518 AA;  57745 MW;  E6FC2EF76DD814DF CRC64;

  Query Match             4.8%;  Score 308;  DB 1;  Length 518;
  Best Local Similarity   27.1%;  Pred. No. 8.6e-11;
  Matches  120;  Conservative   56;  Mismatches  162;  Indels  104;  Gaps   15;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKS-TPLHFAAGFGRKDVVEYLLQNGAN---VQA 160
              | | |:|: :::|:  :|  | | |  :   |:| ||     : :: |:   ::   ::
Db         17 AAREGNVKVLRKLL--KKGRSVDVADNRGWMPIHEAAYHNSVECLQMLINADSSENYIKM 74

Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQH 220
              :   |   || | | || ::| :||  ||||||      |||  |   |:|||  :||||
Db         75 KTFEGFCALHLAASQGHWKIVQILLEAGADPNATTLEETTPLFLAVENGQIDVLRLLLQH 134

Qy        221 GAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVN-CH 279
              ||                                                    ||| |
Db        135 GA-------------------------------------------------NVNGSH 142

Qy        280 ASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVK 339
              :  |  |  || |:     :|::|||: ||:   :|   : ||  |  ||  |   :|:
Db        143 SMCGWNS--LHQASFQENAEIIKLLLRKGANKECQDDFGITPLFVAAQYGKLESLSILIS 200

Qy        340 HGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKER 399
              || ||   | : |||  || :   :   |||| |||| |  |:  |          ||
Db        201 SGANVNCQALDKATPLFIAAQEGHTKCVELLLSSGADPDLY-CNEDSW-------QLPIH 252

Qy        400 LAYEFKGHS-----LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRK 454
              | :  ||:     |:     |  | : |   :     |     ||  |       : |
Db        253 AAAQM-GHTKILDLLIPLTNRACDTGLNKVSPVYSAVF---GGHEDCLEILLRNGYSPDA 308

Qy        455 QICELLLRKGANINEKTKEFLTPLHVASEK--AHNDVVEVVVKHEAKVNALDNLGQTSLH 512
              | | :               | :|: :| :|      :| :::|: |::| |
Db        309 QACLVF------------GFSSPVCMAFQKDCEFFGIVNILLKYGAQINEL--------- 347

Qy        513 RAAYC---GHLQTCRLLLSYGC 531
                |||         |  |  ||
Db        348 HLAYCLKYEKFSIFRYFLRKGC 369
```

```
RESULT 35
PPGB_BOVIN
ID   PPGB_BOVIN     STANDARD;      PRT;   568 AA.
AC   Q95N27;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
DE   inhibited membrane-associated protein) (bTIMAP) (CAAX box protein
DE   TIMAP).
GN   PPP1R16B.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22049828; PubMed=12055102;
RA   Cao W., Mattagajasingh S.N., Xu H., Kim K., Fierlbeck W., Deng J.,
RA   Lowenstein C.J., Ballermann B.J.;
RT   "TIMAP, a novel CAAX box protein regulated by TGF-beta1 and expressed
RT   in endothelial cells.";
RL   Am. J. Physiol. 283:C327-C337(2002).
CC   -!- FUNCTION: May be a downstream target for TGF-beta1 signaling
CC       cascade in endothelial cells.
CC   -!- SUBUNIT: Binds PP1 (Probable).
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (Potential).
CC   -!- INDUCTION: Inhibited by TGF-beta1.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF362909; AAK52795.1; -.
DR   GO; GO:0005886; C:plasma membrane; ISS.
DR   GO; GO:0019903; F:protein phosphatase binding; ISS.
DR   GO; GO:0007165; P:signal transduction; ISS.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   DOMAIN       15     55       COILED COIL (POTENTIAL).
FT   REPEAT      100    129       ANK 1.
FT   REPEAT      133    162       ANK 2.
FT   REPEAT      228    257       ANK 3.
FT   REPEAT      261    290       ANK 4.
FT   REPEAT      531    560       ANK 5.
FT   LIPID       565    565       S-farnesyl cysteine (Potential).
SQ   SEQUENCE   568 AA;  63643 MW;  61C10899462B0334 CRC64;

  Query Match             4.8%;  Score 307.5;  DB 1;  Length 568;
  Best Local Similarity   28.3%;  Pred. No. 1.1e-10;
  Matches   97;  Conservative   61;  Mismatches  138;  Indels   47;  Gaps   13;

Qy        110 DVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFG------RKDV--VEYLLQNGANVQAR 161
              |::  ||    :    | | ||:   : | |           | |   | | |:|  :
Db         43 DLQHRKR----KHERKRSTGGRRK-KVSFEASVALLEASLRNDAEEVRYFLKNKVSPDLC 97

Qy        162 DDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHG 221
              :: ||  ||  |     |:| ||| |||: ||:||  :|||| ||  | |::  :|:|:|
Db         98 NEDGLTALHQCCIDNFEEIVKLLLSHGANVNAKDNELWTPLHAAATCGHINLVKILVQYG 157

Qy        222 AEPTIRNTDGRTALDLA--DPSAKAVLT-GEYK--KDELLESARSGNEEKMMALLTPLNV 276
              |:    |:||    ||   :|:   : |   |:    | :   |:  |::|::     ::
```

```
Db        158 ADLLAVNSDGNMPYDLCEDEPTLDVIETCMAYQGITQEKINEMRAAPEQQMIS-----DI 212

Qy        277 NCHAS--------DGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSY 328

Db        213 HCMIAAGQDLDWVDAQGATLLHIAGANGYLRAAELLLDHGVRVDVKDWDGWEPLHAAAFW 272

Qy        329 GHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN-CHNKSA 387

Db        273 GQMQMAELLVSHGASLSARTSMDEMPI-DLCEEEEFKVLLLELKHKHDVIMKSQLRHKSS 331

Qy        388 ID----------LAPTPQLKER---LAYEFKGHSLLQAAREA 416

Db        332 LSRRTSSAGSRGKVVRRASLSDRTNLYRKEYEGEAILWQQRSA 374


RESULT 36
RN5A_MOUSE
ID   RN5A_MOUSE     STANDARD;      PRT;   735 AA.
AC   Q05921; Q9ERU7;
DT   01-NOV-1995 (Rel. 32, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN   RNASEL OR RNS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3H/An; TISSUE=Adipose tissue;
RX   MEDLINE=20515260; PubMed=11063255;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes: mediators
RT   of interferon action.";
RL   Mamm. Genome 11:989-992(2000).
RN   [2]
RP   SEQUENCE OF 1-679 FROM N.A.
RX   MEDLINE=93201598; PubMed=7680958;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAase: a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:753-765(1993).
RN   [3]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11585831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";
RL   J. Biol. Chem. 276:48473-48482(2001).
RN   [4]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13354-13354(2002).
CC   -!- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of mRNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 3' of UpNp dimers, with preference for
CC       UU and UA sequences, to sets of discrete products ranging from
CC       between 4 and 22 nucleotides in length.
CC   -!- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC       either manganese or magnesium and ATP.
CC   -!- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC       linked oligoadenylates) the homodimerization and subsequent
CC       activation occurs. Inhibited by RNase L inhibitor.
CC   -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC   -!- TISSUE SPECIFICITY: Expressed in spleen, thymus, lung, testis,
CC       kidney, liver and heart.
CC   -!- INDUCTION: By interferons.
CC   -!- DOMAIN: A single active nuclease domain in a dimer is
CC       sufficient for ribonuclease activity.
CC   -!- DOMAIN: The kinase domain allows the homodimerization.
CC   -!- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC       constitute the 2-5A binding domain.
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF281045; AAG33708.1; -.
DR   EMBL; L10382; AAA37117.1; -.
DR   PIR; B45771; B45771.
DR   HSSP; P42773; 1IHB.
DR   MGD; MGI:1098272; Rnasel.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR006567; PUG.
DR   Pfam; PF00023; ank; 8.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00580; PUG; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger;
KW   Repeat; ANK repeat.
FT   REPEAT       24     53       ANK 1.
FT   REPEAT       58     87       ANK 2.
FT   REPEAT       91    120       ANK 3.
FT   REPEAT      124    153       ANK 4.
FT   REPEAT      167    197       ANK 5.
FT   REPEAT      201    234       ANK 6.
FT   REPEAT      238    268       ANK 7.
FT   REPEAT      272    301       ANK 8.
FT   REPEAT      303    328       ANK 9.
FT   DOMAIN      364    584       PROTEIN KINASE LIKE.
FT   DOMAIN      585    735       RIBONUCLEASE.
FT   DOMAIN      229    242       2-5A BINDING (P-LOOP) 1.
FT   DOMAIN      253    275       2-5A BINDING (P-LOOP) 2.
FT   ZN_FING     401    436       C6-TYPE (POTENTIAL).
SQ   SEQUENCE   735 AA;  83274 MW;  B6632F4A5B50F711 CRC64;

  Query Match            4.7%;  Score 305;  DB 1;  Length 735;
  Best Local Similarity   27.5%;  Pred. No. 2.2e-10;
  Matches  148;  Conservative   69;  Mismatches  232;  Indels   90;  Gaps   23;

Qy        565 EADRQLLEAAKAGDVETVKKLCT--VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 622

Db         24 EDDSSLIKAVQKGDVVRVQQLLEKGADANACEDTWG--WTPLHNAVQAGRVDIVNLLLSH 81

Qy        623 GADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEIC 682

Db         82 GADPHRRKKNGATPFIIAGIQGDVKLLEILLSCGADVNECDENGFTAFMEAAERGNAEAL 141

Qy        683 KLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDA-ALLDAAKKGCL--ARVKKLS 739

Db        142 RFLFAKGAN----------VNLRRQTTKDKRRLKQGGATALMSAAEKGHLEVLRILLND 190

Qy        740 SPDNVNCRDTQGRHS---TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 796

Db        191 MKAEVDARDNMGRNALIRTLLNWDC-ENVEEITSILIQHGADVNVRGERGKTPLIAAVER 249

Qy        797 GHVDVAALLIKYNAC-VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 855
```

```
Db        250 KHTGLVQMLLSREGINIDARDNEGKTALLIAVDKQLKEIVQLLLEKGAD---------- 298

Qy        856 LDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSL 915

Db        299 ----KCDDL-VWIARRNHDYHLVKLLLPYVANPDTDPPA-------GDWSPHSSRWGTAL 346

Qy        916 DNLSG----SFSELSSVVSS----SGT-EGASSL---EKKEVPGVDF--------SITQF 955

Db        347 KSLHSMTRPMIGKLKIFIHDDYKIAGTSEGAVYLGIYDNREVAVKVFRENSPRGCKEVSC 406

Qy        956 VRNLGLEH--LMDIFERE------------QITLDVLVEMGHKELKEIGINAYGHR--H 998

Db        407 LRDCG-DHSNLVAFYGREDDKGCLYVCVSLCEWTLEEFLRLPREEPVENGEDKFAHSILL 465

Qy        999 KLIKGVERL-ISG--QQGLNPYLTLNTSGSGTILIDLSPDDK---EFQSVEEEMQSTVR 1051

Db        466 SIFEGVQKLHLHGYSHQDLQPQNILIDSKKAVRLADFDQSIRWMGESQMVRRDLEDLGR 524
```

RESULT 37
PPGB_HUMAN
ID PPGB_HUMAN STANDARD; PRT; 567 AA.
AC Q96T49; O94912; Q9NQG4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
DE inhibited membrane-associated protein) (hTIMAP) (CAAX box protein
DE TIMAP) (Ankyrin repeat domain protein 4).
GN PPP1R16B OR ANKRD4 OR KIAA0823.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22049828; PubMed=12055102;
RA Cao W., Mattagajasingh S.N., Xu H., Kim K., Fierlbeck W., Deng J.,
RA Lowenstein C.J., Ballermann B.J.;
RT "TIMAP, a novel CAAX box protein regulated by TGF-beta1 and expressed
RT in endothelial cells.";
RL Am. J. Physiol. 283:C327-C337(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- FUNCTION: May be a downstream target for TGF-beta1 signaling
CC cascade in endothelial cells.
CC -!- SUBUNIT: Binds PP1 (Probable).
CC -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (Potential).
CC -!- TISSUE SPECIFICITY: Highly expressed in vascular endothelium, CNS,
CC lung, spleen, kidney and testis.
CC -!- INDUCTION: Inhibited by TGF-beta1.
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; AF362910; AAK52796.1; -.
DR EMBL; AB020630; BAA74846.2; ALT_INIT.
DR EMBL; AL121889; CAB98284.1; -.
DR EMBL; AL031657; -; NOT_ANNOTATED_CDS.
DR Genew; HGNC:15850; PPP1R16B.
DR GO; GO:0005886; C:plasma membrane; NAS.
DR GO; GO:0019903; F:protein phosphatase binding; NAS.
DR GO; GO:0007165; P:signal transduction; NAS.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 5.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT DOMAIN 15 55 COILED COIL (POTENTIAL).
FT REPEAT 100 129 ANK 1.
FT REPEAT 133 162 ANK 2.
FT REPEAT 228 257 ANK 3.
FT REPEAT 261 290 ANK 4.
FT REPEAT 530 559 ANK 5.
FT LIPID 564 564 S-farnesyl cysteine (Potential).
SQ SEQUENCE 567 AA; 63551 MW; 8418CE26D0035CF1 CRC64;

Query Match 4.7%; Score 304; DB 1; Length 567;
Best Local Similarity 30.9%; Pred. No. 1.7e-10;
Matches 89; Conservative 48; Mismatches 119; Indels 32; Gaps 10;

```
Qy        110 DVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFG------RKDV--VEYLLQNGANVQAR 161

Db         43 DLQHRKR----KHERKRSTGGRRK-KVSPEASVALLEASLRNDAEEVRYFLKNKVSPDLC 97

Qy        162 DDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHG 221

Db         98 NEDGLTALHQCCIDNFEEIVKLLLSHGANVNAKDNELWTPLHAAATCGHINLVKILVQYG 157

Qy        222 AEPTIRNTDGRTALDLA--DPSAKAVLT-GEYK--KDELLESARSGNEEKMMALLTPLNV 276

Db        158 ADLLAVNSDGNMPYDLCEDEPTLDVIETCMAYQGITQEKINEMRVAPEQQMIA-----DI 212

Qy        277 NCHAS--------DGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSY 328

Db        213 HCMIAAGQDLDWIDAQGATLLHIAGANGYLRAAELLLDHGVRVDVKDWDGWEPLHAAAFW 272

Qy        329 GHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGAD 376

Db        273 GQMQMAELLVSHGASLSARTSMDEMPI-DLCEEEEFKVLLLELKHKHD 319
```

```
RESULT 38
AS10_HUMAN
ID   AS10_HUMAN     STANDARD;      PRT;   429 AA.
AC   Q8WXI3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 10 (ASB-10).
GN   ASB10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kile B.T., Hilton D.J., Nicola N.A.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF417920; AAL59159.1; -.
DR   Genew; HGNC:17185; ASB10.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 7.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 7.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT      115    144       ANK 1.
FT   REPEAT      147    176       ANK 2.
FT   REPEAT      180    209       ANK 3.
FT   REPEAT      214    243       ANK 4.
FT   REPEAT      247    289       ANK 5.
FT   REPEAT      293    322       ANK 6.
FT   REPEAT      326    361       ANK 7.
FT   DOMAIN      374    426       SOCS BOX.
SQ   SEQUENCE   429 AA;  46609 MW;  56019889B7BD50DE CRC64;

  Query Match             4.7%;  Score 303;  DB 1;  Length 429;
  Best Local Similarity   30.1%;  Pred. No. 1.3e-10;
  Matches  110;  Conservative   41;  Mismatches  156;  Indels   58;  Gaps   12;

Qy         38 ARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA------GGGAACAS-- 89
              ||  :   | |     | :||| ::          |  ||   |        |   | |
Db         25 ARLVEKPSRGSEEHLKSGPGPIV-----------TRTASGPALAFWQAVLAGDVGCVSRI 73

Qy         90 AAAEAVEPAARELFEACRNGDVERVKRL---VTPEKVNSRDTAGRKSTPLHFAAGFGRKD 146
                |  :   |    :|:     | || :     :  :: |      :|||| ||   |  :
Db         74 LADSSTGLAPDSVFD---TSDPERWRDFRFNIRALRLWSLTYEEELTTPLHVAASRGHTE 130

Qy        147 VVEYLLQNGANVQARDD---GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLH 203
              |:  ||:     :|| |   ||   || ||: ||   |::||  |||||  |      |||
Db        131 VLRLLLRR----RARPDSAPGGRTALHEACAAGHTACVHVLLVAGADPNIADQDGKRPLH 186

Qy        204 EAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG- 262
                | ::   :||: ||        |||:  :   |    |    |  :  :||   | |
Db        187 LCRGPGTLECAELLLRFGA-----RVDGRSEEEEETPLHVAARLGHVELADLL--LRRGA 239

Qy        263 -----NEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
                    | |    ||   :| | :    ::|          :::  |||  |||  | |:

Db        240 CPDARNAEGWTPLLAACDVRCQSITDAEATTARC------LQLCSLLLSAGADADAADQD 293

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH-------EAASKNRVEVCSLL 370
                 ||| ||   ||  | |||:  |    | ||    ||||         | :::   |   |
Db        294 KQRPLHLACRRGHAAVVELLLSCGVSANTMDYGGHTPLHCALQGPAAALAQSPEHVVRAL 353

Qy        371 LSYGA 375
                |::||
Db        354 LNHGA 358


RESULT 39
ANR6_HUMAN
ID   ANR6_HUMAN     STANDARD;      PRT;   692 AA.
AC   Q9Y2G4; Q9NU24; Q9UFQ9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat domain protein 6.
GN   ANKRD6 OR KIAA0957.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [2]
RP   SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).
RC   TISSUE=Testis;
RA   Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 140-692 FROM N.A. (ISOFORM 1).
RA   Tracey A.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9Y2G4-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9Y2G4-2; Sequence=VSP_000270;
CC         Note=No experimental confirmation available;
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB023174; BAA76801.1; -.
DR   EMBL; AL117504; CAB55968.1; -.
DR   EMBL; AL096678; CAB86658.1; -.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:17280; ANKRD6.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Repeat; ANK repeat; Alternative splicing.
FT   REPEAT        9     38       ANK 1.
```

```
FT   REPEAT       41     70       ANK 2.
FT   REPEAT       74    103       ANK 3.
FT   REPEAT      107    136       ANK 4.
FT   REPEAT      140    169       ANK 5.
FT   REPEAT      173    202       ANK 6.
FT   REPEAT      206    235       ANK 7.
FT   REPEAT      239    268       ANK 8.
FT   VARSPLIC    264    264       Q -> QVLRFSRGRSLRKKRERLKEERRAQSVPRDEVAQSK
FT                                (in isoform 2).
FT                                /FTId=VSP_000270.
FT   CONFLICT     87     87       N -> I (IN REF. 2).
SQ   SEQUENCE   692 AA;  75722 MW;  D32115487A80839F CRC64;

  Query Match             4.6%;  Score 300.5;  DB 1;  Length 692;
  Best Local Similarity   24.7%;  Pred. No. 3.7e-10;
  Matches  114;  Conservative   69;  Mismatches  195;  Indels   83;  Gaps   10;

Qy         93 EAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLL 152
              :||   :  |  |   |  | |:|:      :    ||  |||| ||  |   ||: ||
Db          5 DAVAALSERLLVAAYKGQTENVVQLINKGARVAVTKHGR--TPLHLAANKGHLPVVQILL 62

Qy        153 QNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKID 212
              : | ::  :|||    || |    |: |::  |:  |   : :|    | ||||:  |
Db         63 KAGCDLDVQDDGDQTALHRATVVGNTEIIAALIHEGCALDRQDKDGNTALHEASWHGFSQ 122

Qy        213 VCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 272
                :|:: ||    :|  | ||| ||                  ::: | :   ::   :
Db        123 SAKLLVKAGANVLAKNKAGNTALHLA-----------------CQNSHSQSTRVLLLAGS 165

Qy        273 PLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYE 332
              ::  :| |    | ||:|| || : |::|||     || |::    || | :  | :
Db        166 RADLKNNAGD----TCLHVAARYNHLSIIRLLLTAFCSVHEKNQAGDTALHVAAALNHKK 221

Qy        333 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 392
              | ::|:: ||    ::    |||  |   |  ||  ||           : ::
Db        222 VAKILLEAGADTTIVNNAGQTPLETARYHNNPEVALLLTK----------APQGSVSAGD 271

Qy        393 TPQLKERLA-YEFKGHSLLQAAREADVTRIKKHLSLEMVN-FKHPQT-------HETALH 443
              ||  :: :|  |       | |: |       ::  |   |: |  |          |: ||
Db        272 TPSSEQAVARKEEAREEFLSASPEPRAKDDRRRKSRPKVSAFSDPTPPADQQPGHQKNLH 331

Qy        444 CAAASPYPKRKQ-----------------------------------ICELLLRKGAN 466
                |  :||::                                     || |: |  |
Db        332 ---AHNHPKKRNRHRCSSPPPPHEFRAYQLYTLYRGKDGKVMQAPINGCRCEPLINKLEN 388

Qy        467 INEKT-KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNL 506
                | | :|   |   :| :  : ::   | : ::  || |
Db        389 QLEATVEEIKAELGSVQDKMNTKLGQMENKTQHQMRVLDKL 429


RESULT 40
BAT8_HUMAN
ID   BAT8_HUMAN     STANDARD;      PRT;  1210 AA.
AC   Q96KQ7; Q14349; Q96MH5; Q96QD0; Q9UQL8; Q9Y331;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Histone-lysine N-methyltransferase, H3 lysine-9 specific 3
DE   (EC 2.1.1.43) (Histone H3-K9 methyltransferase 3) (H3-K9-HMTase 3)
DE   (HLA-B associated transcript 8) (G9a) (NG36).
GN   BAT8 OR G9A OR NG36.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rowen L., Qin S., Madan A., Dickhoff R., Dors M., Madan A., Hicks P.,
RA   Loretz C., Ratcliffe A., Abbasi N., Shaffer T., Hood L.;
RT   "Sequence of the human major histocompatibility complex class III
RT   region.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Hirakawa M., Yamaguchi H., Imai K., Shimada J., Shiina S., Tamiya G.,
RA   Oka A., Inoko H.;
RT   "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING (ISOFORM 2), AND
RP   TISSUE SPECIFICITY.
RX   MEDLINE=21564388; PubMed=11707778;
RA   Brown S.E., Campbell R.D., Sanderson C.M.;
RT   "Novel NG36/G9a gene products encoded within the human and mouse MHC
RT   class III regions.";
RL   Mamm. Genome 12:916-924(2001).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RC   TISSUE=Salivary gland;
RA   Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA   Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 5-1210 FROM N.A. (ISOFORM 1).
RC   TISSUE=Muscle, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE OF 195-1210 FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Histiocytic lymphoma;
RX   MEDLINE=93207535; PubMed=8457211;
RA   Milner C.M., Campbell R.D.;
RT   "The G9a gene in the human major histocompatibility complex encodes a
RT   novel protein containing ankyrin-like repeats.";
RL   Biochem. J. 290:811-818(1993).
RN   [7]
RP   ENZYMATIC ACTIVITY, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21326082; PubMed=11316813;
RA   Tachibana M., Sugimoto K., Fukushima T., Shinkai Y.;
RT   "Set domain-containing protein, G9a, is a novel lysine-preferring
RT   mammalian histone methyltransferase with hyperactivity and specific
RT   selectivity to lysines 9 and 27 of histone H3.";
RL   J. Biol. Chem. 276:25309-25317(2001).
RN   [8]
RP   IDENTIFICATION IN COMPLEX WITH E2F6; TFDP1; MAX; MGA; EUHMTASE1; CBX3;
RP   RING1; RNF2; MBLR; L3MBTL2 AND YAF2.
RX   MEDLINE=21999559; PubMed=12004135;
RA   Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT   "A complex with chromatin modifiers that occupies E2F- and
RT   Myc-responsive genes in G0 cells.";
RL   Science 296:1132-1136(2002).
```

```
CC   -!- FUNCTION: Histone methyltransferase. Preferentially methylates
CC       Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro). H3 Lys-9
CC       methylation represents a specific tag for epigenetic
CC       transcriptional repression by recruiting HP1 proteins to
CC       methylated histones. Probably targeted to histone H3 by different
CC       DNA-binding proteins like E2F6, MGA, MAX and/or DP1. Also
CC       methylates histone H1 (By similarity).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC       S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC   -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC       E2F6, MGA, MAX, TFDP1, CBX3, BAT8, EUHMTASE1, RING1, RNF2, MBLR,
CC       L3MBTL2 and YAF2.
CC   -!- SUBCELLULAR LOCATION: Nuclear; associates with euchromatic
CC       regions. Does not associate with heterochromatin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=Q96KQ7-1; Sequence=Displayed;
CC       Name=2; Synonyms=NG36G9a-SPI;
CC         IsoId=Q96KQ7-2; Sequence=VSP_002211;
CC       Name=3; Synonyms=NG36;
CC         IsoId=Q96KQ7-3; Sequence=VSP_002212, VSP_002213;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues examined, with high
CC       levels in fetal liver, thymus, lymph node, spleen and peripheral
CC       blood leukocytes and lower level in bone marrow.
CC   -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC       family. Suvar3-9 subfamily.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- SIMILARITY: Contains 1 pre-SET domain.
CC   -!- SIMILARITY: Contains 1 SET domain.
CC   -!- SIMILARITY: Contains 1 post-SET domain.
CC   -!- CAUTION: NG36 and G9a were originally thought to derive from two
CC       separate genes. Ref.3 shows that all G9A transcripts also contain
CC       the in frame coding sequence of NG36.
CC   -!- CAUTION: Ref.1 and Ref.2 sequences differ from that shown due to
CC       erroneous gene model prediction.
CC   -!- CAUTION: It is not known whether Met-1 or Met-21 is the initiator
CC       methionine.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF134726; AAD21811.1; ALT_SEQ.
DR   EMBL; AF134726; AAD21812.1; ALT_SEQ.
DR   EMBL; AP000502; BAB63294.1; ALT_SEQ.
DR   EMBL; AP000502; BAB63295.1; ALT_SEQ.
DR   EMBL; AJ315532; CAC86666.1; -.
DR   EMBL; AK056936; BAB71314.1; -.
DR   EMBL; BC002686; AAH02686.1; ALT_INIT.
DR   EMBL; BC018718; AAH18718.1; ALT_INIT.
DR   EMBL; BC020970; AAH20970.1; ALT_INIT.
DR   EMBL; X69838; CAA49491.1; ALT_INIT.
DR   Genew; HGNC:14129; BAT8.
DR   MIM; 604599; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR003616; PostSET.
DR   InterPro; IPR007728; Pre-SET.
DR   InterPro; IPR001214; SET.
DR   InterPro; IPR003606; Zn2-binding.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF05033; Pre-SET; 1.
DR   Pfam; PF00856; SET; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00468; PreSET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50868; POST_SET; FALSE_NEG.
DR   PROSITE; PS50867; PRE_SET; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW   ANK repeat; Repeat; Alternative splicing.
FT   DOMAIN        2     13       POLY-ALA.
FT   DOMAIN      160    163       POLY-ALA.
FT   DOMAIN      300    326       POLY-GLU.
FT   REPEAT      649    678       ANK 1.
FT   REPEAT      684    713       ANK 2.
FT   REPEAT      717    746       ANK 3.
FT   REPEAT      750    780       ANK 4.
FT   REPEAT      784    813       ANK 5.
FT   REPEAT      817    846       ANK 6.
FT   REPEAT      850    879       ANK 7.
FT   DOMAIN      972   1035       PRE-SET.
FT   DOMAIN     1037   1159       SET.
FT   DOMAIN     1164   1180       POST-SET.
FT   VARSPLIC    373    406       Missing (in isoform 2).
FT                                /FTId=VSP_002211.
FT   VARSPLIC    195    202       PPVPEKRP -> VSGMGEMG (in isoform 3).
FT                                /FTId=VSP_002212.
FT   VARSPLIC    203   1210       Missing (in isoform 3).
FT                                /FTId=VSP_002213.
FT   CONFLICT      5      5       A -> Q (IN REF. 5; AAH18718).
FT   CONFLICT     13     13       A -> G (IN REF. 5; AAH20970).
FT   CONFLICT     55     55       N -> T (IN REF. 1 AND 4).
FT   CONFLICT    168    169       PP -> TR (IN REF. 2; AAH02686).
FT   CONFLICT    178    178       P -> S (IN REF. 3).
FT   CONFLICT    985    985       C -> R (IN REF. 3).
FT   CONFLICT    994    994       C -> R (IN REF. 6).
SQ   SEQUENCE   1210 AA;  132383 MW;  E8508923CA66911A CRC64;

  Query Match             4.6%;  Score 297;  DB 1;  Length 1210;
  Best Local Similarity   23.4%;  Pred. No. 1.3e-09;
  Matches  143;  Conservative   93;  Mismatches  227;  Indels  148;  Gaps   22;

Qy         83 GGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGF 142
              ||  | ||    : |    | :|    : || |:|    : : |         |: :
Db        615 GG--CLSAVGLPLGPGREALEKALVIQESERRKKL----RFHPRQ--------LYLSVKQ 660

Qy        143 GR-KDVVEYLLQN-GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              |  : |:  || |   | |:        ||| |   |  |: ::||: ||: || |    |
Db        661 GELQKVILMLLDNLDPNFQSDQQSKRTPLHAAAQKGSVEICHVLLQAGANINAVDKQQRT 720

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              || || :  ::|   ::| |     :  || | |                      :|:
Db        721 PLMEAVVNNHLEVARYMVQRGGCVYSKEEDGSTCLH-------------------HAAK 760

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
               || | :  ||:   |: :| |    ||:  || :  ::::::||  ||||   | : :
Db        761 IGNLEMVSLLLSTGQVDVNAQDSGGWTPIIWAAEHKHIEVIRMLLTRGADVTLTDNEENI 820

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              || |    |   : |:|:      ::|::    |||| || ::  :   | || ||:| |
Db        821 CLHWASFTGSAAIAEVLLNARCDLHAVNYHGDTPLHIAARESYHDCVLLFLSRGANPELR 880

Qy        381 NCHNKSAIDLAP-------TPQLKERLAYEFKGHSLLQAARE--ADVTRIKKHLSLEMVN 431
              |    :| || |          ||  :|      |:  ::  :    || |  ::: :  ||
Db        881 NKEGDTAWDLTPERSDVWFALQLNRKLRLGV-GNRAIRTEKIICRDVARGYENVPIPCVN 939

Qy        432 ----------FKH-PQTHETAL----------HCAAASPYPKRKQIC-ELLLR----KG 464
                       :|:  :  ||:            ||           :| :| :|    |
Db        940 GVDGEPCPEDYKYISENCETSTMNIDRNITHLQHCTCVDDCSSSNCLCGQLSIRCWYDKD 999

Qy        465 ANINEKTKEFLTPLHVASEKA-------HNDVV-------------------------- 490
                : ::  :   ||      :|       | ||
Db       1000 GRLLQEFNKIEPPLIFECNQACSCWRNCKNRVVQSGIKVRLQLYRTAKMGWGVRALQTIP 1059

Qy        491 ----------EVVVKHEAKVNA-------LDNL-GQTSLHRAAYCGHLQTCRLLLSYGCD 532
                       |::  || |         ||| |:      | | |::      ::: ||
Db       1060 QGTFICEYVGELISDAEADVREDDSYLFDLDNKDGEVYCIDARYYGNISR---FINHLCD 1116
```

```
Qy        533 PNIISLQGFTALQMGNENVQ---------QLLQEGISLGNSEADRQLLEAAKAGDVETVK 583
              |||| :: |    | :::::         : :: |  ||     ||         |:::
Db       1117 PNIIPVRVF----MLHQDLRFPRIAFFSSRDIRTGEELGFDYGDR-------FWDIKSKY 1165

Qy        584 KLCTVQSVNCR 594
                |   |  |:
Db       1166 FTCQCGSEKCK 1176


RESULT 41
Y045_METMA
ID   Y045_METMA     STANDARD;      PRT;   360 AA.
AC   Q8Q0U0;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical ANK-repeat protein MM0045.
GN   MM0045.
OS   Methanosarcina mazei (Methanosarcina frisia).
OC   Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC   Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2209;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Goe1 / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX   MEDLINE=22120827; PubMed=12125824;
RA   Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA   Martinez-Arias R., Henne A., Wiezer A., Baeumer S., Jacobi C.,
RA   Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA   Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA   Fritz H.-J., Gottschalk G.;
RT   "The genome of Methanosarcina mazei: evidence for lateral gene
RT   transfer between Bacteria and Archaea.";
RL   J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE013224; AAM29741.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat; Complete proteome.
FT   REPEAT      100    129       ANK 1.
FT   REPEAT      133    162       ANK 2.
FT   REPEAT      166    195       ANK 3.
FT   REPEAT      199    228       ANK 4.
FT   REPEAT      232    261       ANK 5.
FT   REPEAT      265    294       ANK 6.
FT   REPEAT      298    327       ANK 7.
SQ   SEQUENCE   360 AA;  39044 MW;  C20B57832367A592 CRC64;

  Query Match             4.6%;  Score 295.5;  DB 1;  Length 360;
  Best Local Similarity   29.8%;  Pred. No. 2.7e-10;
  Matches   97;  Conservative   56;  Mismatches  142;  Indels   31;  Gaps    7;

Qy        552 QQLLQEGIS-LGNSEADRQLLEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAA 608
              :::| : || |   : ::  |||:| |  | |:||      || :|  |:  | |  ||
Db         52 EKILMDFISNLFGRDKNQSFLEASKQGQTENVEKLLRSNKVDVNYQDAYGK--TALISAA 109

Qy        609 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 668
                    |:  |:: | ::  :|: |   | :|      :: :||||:|| :|  |    |
Db        110 DKGYRDVIGLLIESGPNLDLQDENGNTALISAAKIERGDIIDLLVKNGADLNFQDENGET 169

Qy        669 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD-GDTDIQDLLRGDAALLDAA 727
              |  |   |   |  |:  | |   :: :| | |    |    || :||
Db        170 ALKFAVKVGYKNIADQLIDAGTDLNIQDENGETALICAADRAHRDIAELL---------- 219

Qy        728 KKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGL 787
                    : :  ::| :|  |:  | |  |    :  : | |:  |||:| ||| |
Db        220 ----------IKAGADLNIQDNSGK--TALVAATKIGHKGIVELLVNAGADLNLQDKNGN 267

Qy        788 IPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTL 847
              |  ||  |: |:  |||:  | :|  |:  | |  :|| ||  :  ||:  |||  :
Db        268 TALIYAADRGYRDIVNLLIEGGASLNIPDEAGLTALMFSAQTGRKDIVELLIKAGADINI 327

Qy        848 KNQEGQTPLDL---VSADDVSALLTA 870
              ::: :|  ||   |  :::  |||:
Db        328 EDKNNKTAADLAAEVGFEEIVDLLTS 353


RESULT 42
RN5A_HUMAN
ID   RN5A_HUMAN     STANDARD;      PRT;   741 AA.
AC   Q05823;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN   RNASEL OR RNS4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-240 AND LYS-274.
RC   TISSUE=Kidney;
RX   MEDLINE=93201598; PubMed=7680958;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAase: a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:753-765(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20515260; PubMed=11063255;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes: mediators
RT   of interferon action.";
RL   Mamm. Genome 11:989-992(2000).
RN   [3]
RP   CHARACTERIZATION OF RNASEL ACTIVITY.
RX   MEDLINE=94245737; PubMed=7514601;
RA   Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
RA   Silverman R.H.;
RT   "Intrinsic molecular activities of the interferon-induced 2-5A-
RT   dependent RNase.";
RL   J. Biol. Chem. 269:14153-14158(1994).
RN   [4]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11585831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";
RL   J. Biol. Chem. 276:48473-48482(2001).
RN   [5]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13354-13354(2002).
RN   [6]
RP   REVIEW.
RX   MEDLINE=99073516; PubMed=9856285;
RA   Castelli J., Wood K.A., Youle R.J.;
RT   "The 2-5A system in viral infection and apoptosis.";
RL   Biomed. Pharmacother. 52:386-390(1998).
```

```
RN   [7]
RP   MUTAGENESIS OF LYS-392.
RX   MEDLINE=99081839; PubMed=9862963;
RA   Dong B., Silverman R.H.;
RT   "Alternative function of a protein kinase homology domain in 2',
RT   5'-oligoadenylate dependent RNase L.";
RL   Nucleic Acids Res. 27:439-445(1999).
RN   [8]
RP   MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND
RP   HIS-672.
RX   MEDLINE=21230542; PubMed=11333017;
RA   Dong B., Niwa M., Walter P., Silverman R.H.;
RT   "Basis for regulated RNA cleavage by functional analysis of RNase L
RT   and Irelp.";
RL   RNA 7:361-373(2001).
RN   [9]
RP   VARIANTS SER-59; PHE-406; GLN-462 AND GLU-541.
RX   MEDLINE=21947464; PubMed=11941539;
RA   Roekman A., Ikonen T., Seppaelae E.H., Nupponen N., Autio V.,
RA   Mononen N., Bailey-Wilson J., Trent J., Carpten J., Matikainen M.P.,
RA   Koivisto P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
RT   "Germline alterations of the RNASEL gene, a candidate HPC1 gene at
RT   1q25, in patients and families with prostate cancer.";
RL   Am. J. Hum. Genet. 70:1299-1304(2002).
RN   [10]
RP   VARIANTS SER-59; GLN-462 AND GLU-541.
RX   MEDLINE=21676832; PubMed=11799394;
RA   Carpten J., Nupponen N., Isaacs S., Sood R., Robbins C., Xu J.,
RA   Faruque M., Moses T., Ewing C., Gillanders E., Hu P., Bujnovszky P.,
RA   Makalowska I., Baffoe-Bonnie A., Faith D., Smith J., Stephan D.,
RA   Wiley K., Brownstein M., Gildea D., Kelly B., Jenkins R.,
RA   Hostetter G., Matikainen M., Schleutker J., Klinger K., Connors T.,
RA   Xiang Y., Wang Z., De Marzo A., Papadopoulos N., Kallioniemi O.-P.,
RA   Burk R., Meyers D., Groenberg H., Meltzer P., Silverman R.,
RA   Bailey-Wilson J., Walsh P., Isaacs W., Trent J.;
RT   "Germline mutations in the ribonuclease L gene in families showing
RT   linkage with HPC1.";
RL   Nat. Genet. 30:181-184(2002).
RN   [11]
RP   CHARACTERIZATION OF VARIANTS GLN-462 AND GLU-541.
RX   MEDLINE=22345000; PubMed=12415269;
RA   Casey G., Neville P.J., Plummer S.J., Xiang Y., Krumroy L.M.,
RA   Klein E.A., Catalona W.J., Nupponen N., Carpten J.D., Trent J.M.,
RA   Silverman R.H., Witte J.S.;
RT   "RNASEL Arg462Gln variant is implicated in up to 13% of prostate
RT   cancer cases.";
RL   Nat. Genet. 32:581-583(2002).
CC   -!- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of mRNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 3' of UpNp dimers, with preference for
CC       UU and UA sequences, to sets of discrete products ranging from
CC       between 4 and 22 nucleotides in length.
CC   -!- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC       either manganese or magnesium and ATP.
CC   -!- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC       linked oligoadenylates) the homodimerization and subsequent
CC       activation occurs. Inhibited by RNase L inhibitor.
CC   -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC   -!- TISSUE SPECIFICITY: Highly expressed in spleen and thymus followed
CC       by prostate, testis, uterus, small intestine, colon and peripheral
CC       blood leukocytes.
CC   -!- INDUCTION: By interferons.
CC   -!- DOMAIN: The kinase domain allows the homodimerization.
CC   -!- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC       constitute the 2-5A binding domain.
CC   -!- DISEASE: Defects in RNASEL may be the cause of susceptibility to
CC       hereditary prostate cancer 1 (HPC1) [MIM:601518].
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10381; AAA18032.1; -.
DR   HSSP; P80144; 2MYO.
DR   Genew; HGNC:10050; RNASEL.
DR   MIM; 180435; -.
DR   MIM; 601518; -.
DR   GO; GO:0004521; F:endoribonuclease activity; NAS.
DR   GO; GO:0003723; F:RNA binding; IDA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR006567; PUG.
DR   Pfam; PF00023; ank; 9.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00580; PUG; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger; Repeat;
KW   ANK repeat; Polymorphism.
FT   REPEAT      24     53      ANK 1.
FT   REPEAT      58     87      ANK 2.
FT   REPEAT      91    120      ANK 3.
FT   REPEAT     124    153      ANK 4.
FT   REPEAT     167    197      ANK 5.
FT   REPEAT     201    234      ANK 6.
FT   REPEAT     238    268      ANK 7.
FT   REPEAT     272    301      ANK 8.
FT   REPEAT     303    329      ANK 9.
FT   DOMAIN     365    586      PROTEIN KINASE LIKE.
FT   DOMAIN     587    741      RIBONUCLEASE.
FT   DOMAIN     229    242      2-5A BINDING (P-LOOP) 1.
FT   DOMAIN     253    275      2-5A BINDING (P-LOOP) 2.
FT   ZN_FING    395    444      C6-TYPE (ATYPICAL).
FT   VARIANT     59     59      G -> S.
FT                              /FTId=VAR_013509.
FT   VARIANT    406    406      S -> F.
FT                              /FTId=VAR_013510.
FT   VARIANT    462    462      R -> Q (risk factor for prostate cancer;
FT                              reduced enzymatic activity;
FT                              dbSNP:486907).
FT                              /FTId=VAR_012056.
FT   VARIANT    541    541      D -> E (no change in enzymatic activity;
FT                              dbSNP:627928).
FT                              /FTId=VAR_012057.
FT   MUTAGEN    240    240      K->N: REDUCED 2-5A BINDING ACTIVITY;
FT                              ALMOST COMPLETE LOSS OF 2-5A BINDING
FT                              ACTIVITY; WHEN ASSOCIATED WITH N-274.
FT   MUTAGEN    274    274      K->N: REDUCED 2-5A BINDING ACTIVITY;
FT                              ALMOST COMPLETE LOSS OF 2-5A BINDING
FT                              ACTIVITY; WHEN ASSOCIATED WITH N-240.
FT   MUTAGEN    392    392      K->R: COMPLETE LOSS OF ENZYMATIC ACTIVITY
FT                              AND ENZYME DIMERIZATION. NO CHANGE IN
FT                              BINDING TO 2-5A AND RNA.
FT   MUTAGEN    583    583      H->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN    584    584      P->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN    632    632      W->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN    661    661      D->A: COMPLETE LOSS OF ENZYMATIC
FT                              ACTIVITY.
FT   MUTAGEN    667    667      R->A: COMPLETE LOSS OF ENZYMATIC
FT                              ACTIVITY. NO CHANGE IN 2-5A BINDING AND
FT                              ENZYME DIMERIZATION.
FT   MUTAGEN    672    672      H->A: COMPLETE LOSS OF ENZYMATIC
FT                              ACTIVITY. NO CHANGE IN 2-5A BINDING
```

```
  Query Match             4.5%;  Score 291.5;  DB 1;  Length 741;
  Best Local Similarity   23.5%;  Pred. No. 1.4e-09;
  Matches  149;  Conservative   83;  Mismatches  210;  Indels  191;  Gaps   24;

Qy        146 DVVEYLLQNGANVQ-ARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHE 204
              |:|: ||: ||||    ::||  |||||       ::| |||||||| |       ||
Db         39 DLVQQLLEGGANVNFQEEEGGWTPLHNAVQMSREDIVELLLRHGADPVLRKKNGATPFIL 98

Qy        205 AAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNE 264
              ||| | : :  : |  ||:    :  | ||                    :|
Db         99 AAIAGSVKLLKLFLSKGADVNECDFYGFTA-------------------FME------- 131

Qy        265 EKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK---------- 314
                                       || | :|| :: | : ||:|: :
Db        132 --------------------------AAVYGKVKALKFLYKRGANVNLRRKTKEDQERL 164

Qy        315 DKGDLVPLHNACSYGHYEVTELLV-KHGACVNAMD-LWQFTPLHEAASKNRVEV---CSL 369
               ||    | :|   || || ::|: : || ||| | : :   :|   | :  :|      |
Db        165 RKGGATALMDAAEKGHVEVLKILLDEMGADVNACDNMGRNALIHALLSSDDSDVEAITHL 224

Qy        370 LLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEM 429
              || :|||  :      |        || :   || | |   |:|   |       ::|: :
Db        225 LLDHGADVNVRGERGK-------TPLI---LAVEKKHLGLVQRLLE------QEHIEIND 268

Qy        430 VNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDV 489
               :      :|||  |        | |:| ||| ::||     : :    :  |     : :
Db        269 TD----SDGKTALLLAVEL---KLKKIAELLCKRGA-----STDCGDLVMTARRNYDHSL 316

Qy        490 VEVVVKHEAKVN----ALDNLGQTS--------LHR---------------------- 513
              |:|:: | || :    | |   |:|        |||
Db        317 VKVLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLKFFIDEKYKIADTSE 376

Qy        514 -AAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 572
               | |  :   : :   |: :  : :  : ||   ||          ::   ||: |
Db        377 GGIYLGFYEKQEVAVKTFCEGSPRAQREVSCLQSSREN-----SHLVTFYGSESHR---- 427

Qy        573 AAKAGDVETVKKLCTVQSVNCRDI---EGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAK 629
                 | :     ||     | |:   |  ::    ||   | :| :   :|
Db        428 ----GHLFVCVTLCEQTLEACLDVHRGEDVENEEDEFAR--NVLSSIFKAVQE------- 474

Qy        630 DKGGLVPLHNACSYGHYEV--AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQ 687
                    || :| | | ::    :|:      ::||  |                  ::
Db        475 -------LHLSCGYTHQDLQPQNILIDSKKAAHLADFDK-----------------SIK 509

Qy        688 HGADPTKKNRD----GNTPLDLVKDGDTDIQDL 716
                 || :  ||     |    | :|| |    :||
Db        510 WAGDPQEVKRDLEDLGRLVLYVVKKGSISFEDL 542


RESULT 43
ORP1_HUMAN
ID   ORP1_HUMAN     STANDARD;      PRT;   950 AA.
AC   Q9BXW6; Q9BZF5; Q9NW87;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Oxysterol binding protein-related protein 1 (OSBP-related protein 1)
DE   (ORP-1).
GN   OSBPL1A OR OSBPL1 OR ORP1 OR OSBP8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=21601154; PubMed=11735225;
RA   Jaworski C.J., Moreira E., Li A., Lee R., Rodriguez I.R.;
RT   "A family of 12 human genes containing oxysterol-binding domains.";
RL   Genomics 78:185-196(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=21264638; PubMed=11279184;
RA   Xu Y., Liu Y., Ridgway N.D., McMaster C.R.;
RT   "Novel members of the human oxysterol-binding protein family bind
RT   phospholipids and regulate vesicle transport.";
RL   J. Biol. Chem. 276:18407-18414(2001).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM C).
RX   MEDLINE=21376257; PubMed=11483621;
RA   Lehto M., Laitinen S., Chinetti G., Johansson M., Ehnholm C.,
RA   Staels B., Ikonen E., Olkkonen V.M.;
RT   "The OSBP-related protein family in humans.";
RL   J. Lipid Res. 42:1203-1213(2001).
RN   [4]
RP   SEQUENCE OF 514-790 FROM N.A. (ISOFORM A).
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds phospholipids; exhibits strong binding to
CC       phosphatidic acid and weak binding to phosphatidylinositol 3-
CC       phosphate (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=B; Synonyms=OSBPL1B, OSBP8L;
CC         IsoId=Q9BXW6-1; Sequence=Displayed;
CC       Name=A; Synonyms=OSBPL1A, OSBP8S;
CC         IsoId=Q9BXW6-2; Sequence=VSP_003779;
CC       Name=C;
CC         IsoId=Q9BXW6-3; Sequence=VSP_003780;
CC   -!- SIMILARITY: Belongs to the OSBP family.
CC   -!- SIMILARITY: Contains 3 ANK repeats.
CC   -!- SIMILARITY: Contains 1 PH domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF392449; AAL40662.1; -.
DR   EMBL; AF392450; AAL40663.1; -.
DR   EMBL; AF274714; AAK15154.1; -.
DR   EMBL; AF323726; AAG53407.2; -.
DR   EMBL; AK001079; BAA91496.1; -.
DR   Genew; HGNC:16398; OSBPL1A.
DR   MIM; 606730; -.
DR   GO; GO:0005622; C:intracellular; IC.
DR   GO; GO:0005543; F:phospholipid binding; NAS.
DR   GO; GO:0008203; P:cholesterol metabolism; NAS.
DR   GO; GO:0016192; P:vesicle-mediated transport; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000648; Oxysterol_BP.
DR   InterPro; IPR001849; PH.
DR   Pfam; PF00023; ank; 3.
DR   Pfam; PF01237; Oxysterol_BP; 1.
DR   Pfam; PF00169; PH; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 3.
DR   SMART; SM00233; PH; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 3.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS01013; OSBP; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Lipid transport; Transport; ANK repeat; Repeat; Alternative splicing.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT       80    109       ANK 2.
FT   REPEAT      175    204       ANK 3.
FT   DOMAIN      235    334       PH.
FT   VARSPLIC      1    513       Missing (in isoform A).
```

```
FT                                /FTId=VSP_003779.
FT   VARSPLIC    625    648       FNPLLGETYELVRDDLGFRLISEQ -> L (in isoform
FT                                C).
FT                                /FTId=VSP_003780.
FT   CONFLICT    735    735       T -> A (IN REF. 4).
FT   CONFLICT    810    810       S -> P (IN REF. 1; AAL40663).
FT   CONFLICT    841    841       P -> S (IN REF. 1; AAL40662).
FT   CONFLICT    844    844       A -> G (IN REF. 1; AAL40662).
SQ   SEQUENCE   950 AA;  108469 MW;  CDDA26CA27B65F63 CRC64;

  Query Match             4.5%;  Score 291.5;  DB 1;  Length 950;
  Best Local Similarity   23.9%;  Pred. No. 2e-09;
  Matches  142;  Conservative   91;  Mismatches  221;  Indels  139;  Gaps   23;

Qy        563 NSEADRQLLEAAKAGDVETVKKLC-------TVQSVNCRDIEGRQSTPLHFAAGYNRVSV 615
              |:||::|||  |: |: | |::|          :  :||:   ||  : |
Db          2 NTEAEQQLLHHARNGNAEEVRQLLETMARNEVIADINCK---GRSKSNL----------- 47

Qy        616 VEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAA 675
                          |  ||| || :|| :| : |:| || ||| :    |||| ||
Db           48 ----------------GWTPLHLACYFGHRQVVQDLLKAGAEVNVLNDMGDTPLHRAAF 90

Qy        676 KGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG---------DAALLDA 726
               |: |:  |||:: || |  |  | |  ::      :|: :|            :  || |
Db          91 TGRKELVMLLLEYNADTTIVNGSGQTAKEVTH--AEEIRSMLEAVERTQQRKLEELLLAA 148

Qy        727 AKKGCLARVKKL---SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQD 783
              |::|    :  |    :| :||| |  |   :|||| ||   : :  |  ||: ||| | ::
Db        149 AREGKTTELTALLNRPNPPDVNCSDQLG--NTPLHCAAYRAHKQCALKLLRSGADPNLKN 206

Qy        784 KGGLIPLHNA--ASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAH 841
              |   ||  |  |   |: |   :| | |      |    || ::::   :|  ::| |
Db        207 KNDQKPLDLAQGAEMKHILVGNKVI-YKAL-----KRYEGPLWKSSRFFGWRLFWVVLEH 260

Qy        842 GADPTLKNQ---------EGQTPLDLVSADDVSALLTAAMPPSAL--PSCYKPQVLNGVR 890
              |     : |          :|  |          |: |     | |     |:    ::| |
Db        261 GVLSWYRKQPDAVHNIYRQGCKHL-------TQAVCTVKSTDSCLFFIKCF-DDTIHGFR 312

Qy        891 SP-----GATADALSSGPSSPSSLSAASSLDNLSGSFSE--LSSVVSSSGTEGASSLEKK 943
              |      :  | | :     :  :   | | |:     |  :|:      | | | :::
Db        313 VPKNSLQQSREDWLEAIEEHSAYSTHYCSQDQLTDEEEEDTVSAADLKKSLEKAQSCQQR 372

Qy        944 EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKG 1003
                :|  |: |:: :                |:  ||     |::: : :   |   :
Db        373 ----LDREISNFLKMIK-------------ECDMAKEMLPSFLQKVEVVSEASRETCV-- 413

Qy       1004 VERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRD-------G 1056
                     |  | | |  |      |  : :: : : | :::   || :       |
Db        414 ---------ALTDCLNLFTKQEGVRNFKLEQEQEKNKILSEALETLATEHHELEQSLVKG 464

Qy       1057 GHAGGIFNR---YNIL----------KIQKVCNKKLWERYTH---RRKEVSEE 1093
                  | :    |: |           ::: |  :   |   |    |:  :|||
Db        465 SPPASILSEDEFYDALSDSESERSLSRLEAVTARSFEEEGEHLGSRKHRMSEE 517


RESULT 44
KBF2_HUMAN
ID   KBF2_HUMAN     STANDARD;      PRT;   899 AA.
AC   Q00653; Q04860; Q9BU75; Q9H471;
DT   01-OCT-1993 (Rel. 27, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Nuclear factor NF-kappa-B p100/p49 subunits (H2TF1) (Oncogene Lyt-10)
DE   (Lyt10) [Contains: Nuclear factor NF-kappa-B p52 subunit].
GN   NFKB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Leukemia;
RX   MEDLINE=91343004; PubMed=1876189;
RA   Schmid R.M., Perkins N.D., Duckett C.S., Andrews P.C., Nabel G.J.;
RT   "Cloning of an NF-kappa B subunit which stimulates HIV transcription
RT   in synergy with p65.";
RL   Nature 352:733-736(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92123193; PubMed=1531086;
RA   Bours V., Burd P.R., Brown K., Villalobos J., Park S., Ryseck R.P.,
RA   Bravo R., Kelly K., Siebenlist U.;
RT   "A novel mitogen-inducible gene product related to p50/p105-NF-kappa
RT   B participates in transactivation through a kappa B site.";
RL   Mol. Cell. Biol. 12:685-695(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM P80HT).
RX   MEDLINE=94309910; PubMed=8036016;
RA   Thakur S., Lin H.C., Tseng W.T., Kumar S., Bravo R., Foss F.,
RA   Gelinas C., Rabson A.B.;
RT   "Rearrangement and altered expression of the NFKB-2 gene in human
RT   cutaneous T-lymphoma cells.";
RL   Oncogene 9:2335-2344(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Wilson S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE OF 1-220 FROM N.A.
RX   MEDLINE=95059001; PubMed=7969113;
RA   Liptay S., Schmid R.M., Nabel E.G., Nabel G.J.;
RT   "Transcriptional regulation of NF-kappa B2: evidence for kappa B-
RT   mediated positive and negative autoregulation.";
RL   Mol. Cell. Biol. 14:7695-7703(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 37-327.
RX   MEDLINE=98046037; PubMed=9384586;
RA   Cramer P., Larson C.J., Verdine G.L., Mueller C.W.;
RT   "Structure of the human NF-kappaB p52 homodimer-DNA complex at 2.1-A
RT   resolution.";
RL   EMBO J. 16:7078-7090(1997).
CC   -!- FUNCTION: P100 is the precursor of the p52 subunit of the nuclear
CC       factor NF-kappa-B, which binds to the kappa-B consensus sequence
CC       5'-GGRNNYYCC-3', located in the enhancer region of genes involved
CC       in immune response and acute phase reactions. The precursor
CC       protein itself does not bind to DNA. Isoform p49 is a subunit of
CC       the NF-kappa-B protein complex, which stimulates the HIV enhancer
CC       in synergy with p65.
CC   -!- SUBUNIT: Active NF-kappa-B is a heterodimer of an about 52 kDa
CC       DNA-binding subunit and the weak DNA-binding subunit p65. Two
CC       heterodimers might form a labile tetramer.
CC   -!- SUBCELLULAR LOCATION: Nuclear, but also found in the cytoplasm in
```

```
CC       an inactive form complexed to an inhibitor (I-kappa-B).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=p100;
CC         IsoId=Q00653-1; Sequence=Displayed;
CC       Name=p49;
CC         IsoId=Q00653-2; Sequence=VSP_005585, VSP_005586;
CC   -!- DOMAIN: The C-terminus of p100 might be involved in cytoplasmic
CC       retention, inhibition of DNA-binding by p52 homodimers, and/or
CC       transcription activation (By similarity).
CC   -!- DISEASE: In cutaneous T-cell leukemia (CTCL) cell line HUT 78, a
CC       rearranged NFKB2 gene is found which encodes for a truncated 80
CC       kDa protein (p80HT).
CC   -!- SIMILARITY: Belongs to the Rel/Dorsal family.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC       WWW="http://www.infobiogen.fr/services/chromcancer/Genes/NFKB2ID362.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61498; CAA43715.1; -.
DR   EMBL; X61499; CAA43716.1; ALT_SEQ.
DR   EMBL; S76638; AAB21124.1; -.
DR   EMBL; U09609; AAA21462.1; -.
DR   EMBL; AL121928; CAC08399.1; -.
DR   EMBL; BC002844; AAH02844.1; -.
DR   EMBL; U20816; AAA68171.1; -.
DR   PIR; A42024; A42024.
DR   PIR; S17233; S17233.
DR   PDB; 1A3Q; 27-MAY-98.
DR   TRANSFAC; T01927; -.
DR   TRANSFAC; T01930; -.
DR   Genew; HGNC:7795; NFKB2.
DR   MIM; 164012; -.
DR   GO; GO:0003713; F:transcription co-activator activity; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR002909; IPT_TIG.
DR   InterPro; IPR000451; NF_Rel_dor.
DR   InterPro; IPR008967; P53-like.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00554; RHD; 1.
DR   Pfam; PF01833; TIG; 1.
DR   PRINTS; PR00057; NFKBTNSCPFCT.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00429; IPT; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR   PROSITE; PS01204; REL_1; 1.
DR   PROSITE; PS50254; REL_2; 1.
KW   DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW   Repeat; ANK repeat; Phosphorylation; Alternative splicing;
KW   Disease mutation; Proto-oncogene; 3D-structure.
FT   DOMAIN       38    343       REL-LIKE (RHD).
FT   REPEAT      487    519       ANK 1.
FT   REPEAT      526    555       ANK 2.
FT   REPEAT      559    591       ANK 3.
FT   REPEAT      599    628       ANK 4.
FT   REPEAT      633    663       ANK 5.
FT   REPEAT      667    696       ANK 6.
FT   REPEAT      729    758       ANK 7.
FT   DOMAIN      764    851       DEATH.
FT   DOMAIN      337    341       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      350    400       GLY-RICH.
FT   VARSPLIC    374    447       GSLGFFPSSLAYSPYQSGAGPMGCYPGGGGGAQMAATVPSR
FT                                DSGEEAAEPSAPSRTPQCEPQAPEMLQRAREYN -> EGVL
FT                                MEGGVKVREAVEEKNLGEAGRGLHACNPAFGRPRQADYLRS
FT                                GVQDQLGQQRETSSLLKIQTLAGHGGRRL (in isoform
FT                                p49).
FT                                /FTId=VSP_005585.
FT   VARSPLIC    448    899       Missing (In isoform p49).
FT                                /FTId=VSP_005586.
FT   VARIANT     667    669       AGN -> SAS (IN TRUNCATED FORM P80HT).
FT                                /FTId=VAR_006909.
FT   VARIANT     670    899       MISSING (IN TRUNCATED FORM P80HT).
FT                                /FTId=VAR_006910.
FT   CONFLICT    144    144       K -> E (IN REF. 2).
FT   CONFLICT    213    213       I -> T (IN REF. 1, 3 AND 6).
FT   CONFLICT    396    396       G -> R (IN REF. 1 AND 3).
FT   CONFLICT    433    434       EP -> DA (IN REF. 2).
FT   CONFLICT    455    470       QRSARALLDYGVTADA -> HAAPSPTRLLRHRGR (IN
FT                                REF. 1 AND 3).
FT   CONFLICT    470    470       A -> R (IN REF. 2).
FT   CONFLICT    741    741       K -> L (IN REF. 1).
FT   CONFLICT    859    859       T -> TA (IN REF. 4).
FT   CONFLICT    859    859       T -> TT (IN REF. 2).
FT   CONFLICT    875    875       E -> K (IN REF. 2).
FT   CONFLICT    890    890       C -> S (IN REF. 1).
FT   CONFLICT    899    899       H -> TDLLPAPSPLPGPPVQRPHLFQILFNTPHPPLSWDK
FT                                (IN REF. 1).
FT   STRAND       39     44
FT   STRAND       46     46
FT   STRAND       51     51
FT   STRAND       54     55
FT   TURN         56     58
FT   STRAND       67     67
FT   TURN         72     73
FT   STRAND       79     83
FT   TURN         84     84

  Query Match             4.5%;  Score 289;  DB 1;  Length 899;
  Best Local Similarity   26.5%;  Pred. No. 2.6e-09;
  Matches  131;  Conservative   60;  Mismatches  181;  Indels  122;  Gaps   19;

Qy        546 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 605
              :   : : ||  |::     || : | | :       : | | |  |         ||||
Db        453 LAQRSARALLDYGVT-----ADARALLAGQ-------RHLLTAQDEN-------GDTPLH 493

Qy        606 FAAGYNRVSVVE---YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 662
              |  : : ||:|   |:: |  |:                                |||:
Db        494 LAIIHGQTSVIEQIVYVIHHAQDL----------------------------GVVNLT 523

Qy        663 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAA 722
              :    |||| |    |:  :   ||: ||||   :| |:: : |          :|||   |
Db        524 NHLHQTPLHLAVITGQTSVVSFLLRVGADPALLDRHGDSAMHLALRAGAGAPELLR---A 580

Qy        723 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 782
              || :                : |  ::  | :| :  |:||    :  |  : |:  ||:| |
Db        581 LLQSGAP---------AVPQLLHMPDFEGLY--PVHLAVRARSPECLDLLVDSGAEVEAT 629

Qy        783 DK-GGLIPLHNAASYGHVD-VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              :: ||   || |    :  |  |: |  | |||      ||||  ||  |   |  |||
Db        630 ERQGGRTALHLATEMEELGLVTHLVTKLRANVNARTFAGNTPLHLAAGLGYPTLTRLLLK 689

Qy        841 HGADPTLKNQE---------------------------GQTPLDLVSADDVSALLTAA 871
              |||   :|:|                           | |||||  :  |  ||  |
Db        690 AGADIHAENEEPLCPLPSPPTSDSDSDSEGPEKDTRSSFRGHTPLDLTCSTKVKTLLLNA 749

Qy        872 M-----PPSALPSCYKPQVLNGVRSPGATA----DALSSGPSSPSSLSAASSLDNLSGSF 922
                  ||   ||   | :     | | ||    : |  || :  | :  :    |
Db        750 AQNTMEPPLTPPSPAGPGL-----SLGDTALQNLEQLLDGPEAQGSWAELAERLGLRSLV 804

Qy        923 SELSSVVSSSGTEGASSLEKKEVPGVDFS-ITQFVRNLGLEHLMDIFEREQITLDVLVEM 981
                  | ||    | |   |: | | : : : : ::|||  : :  |   | |   ::
```

```
Db        805 DTYRQTTSPSG----SLLRSYELAGGDLAGLLEALSDMGLEEGVRLL-RGPETRD---KL 856

Qy        982 GHKELKEIGINAYG 995
              |:||   :|||
Db        857 PSTEVKE--DSAYG 868


RESULT 45
BAT8_MOUSE
ID   BAT8_MOUSE     STANDARD;      PRT;  1263 AA.
AC   Q9Z148; Q8K4R6; Q8K4R7; Q9Z149;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Histone-lysine N-methyltransferase, H3 lysine-9 specific 3
DE   (EC 2.1.1.43) (Histone H3-K9 methyltransferase 3) (H3-K9-HMTase 3)
DE   (HLA-B associated transcript 8) (G9a) (NG36).
GN   BAT8 OR G9A OR NG36.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129;
RA   Rowen L., Madan A., Qin S., Hall J., Dahl T., James R., Dickhoff R.,
RA   Schaffer T., Ratcliffe A., Abbasi N., Loretz C., Lasky S., Hood L.;
RT   "Sequence of the mouse MHC class III region.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF
RP   1165-ASN--CYS-1168.
RX   MEDLINE=22123403; PubMed=12130538;
RA   Tachibana M., Sugimoto K., Nozaki M., Ueda J., Ohta T., Ohki M.,
RA   Fukuda M., Takeda N., Niida H., Kato H., Shinkai Y.;
RT   "G9a histone methyltransferase plays a dominant role in euchromatic
RT   histone H3 lysine 9 methylation and is essential for early
RT   embryogenesis.";
RL   Genes Dev. 16:1779-1791(2002).
RN   [3]
RP   SEQUENCE OF 960-1263 FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   ALTERNATIVE SPLICING (ISOFORM 2).
RX   MEDLINE=21564388; PubMed=11707778;
RA   Brown S.E., Campbell R.D., Sanderson C.M.;
RT   "Novel NG36/G9a gene products encoded within the human and mouse MHC
RT   class III regions.";
RL   Mamm. Genome 12:916-924(2001).
RN   [5]
RP   ENZYMATIC ACTIVITY, SUBCELLULAR LOCATION, AND MUTAGENESIS OF ARG-1162.
RX   MEDLINE=21326082; PubMed=11316813;
RA   Tachibana M., Sugimoto K., Fukushima T., Shinkai Y.;
RT   "Set domain-containing protein, G9a, is a novel lysine-preferring
RT   mammalian histone methyltransferase with hyperactivity and specific
RT   selectivity to lysines 9 and 27 of histone H3.";
RL   J. Biol. Chem. 276:25309-25317(2001).
CC   -!- FUNCTION: Histone methyltransferase. Preferentially methylates
CC       Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro). H3 Lys-9
CC       methylation represents a specific tag for epigenetic
CC       transcriptional repression by recruiting HP1 proteins to
CC       methylated histones. Also methylates histone H1.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC       S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC   -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC       E2F6, MGA, MAX, TFDP1, CBX3, BAT8, EUHMTASE1, RING1, RNF2, MBLR,
CC       L3MBTL2 and YAF2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear; almost excluded form nucleoli.
CC       Associates with euchromatic regions. Does not associate with
CC       heterochromatin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=G9a-L;
CC         IsoId=Q9Z148-1; Sequence=Displayed;
CC       Name=2; Synonyms=G9a-S;
CC         IsoId=Q9Z148-2; Sequence=VSP_002214, VSP_002215, VSP_002216;
CC   -!- MISCELLANEOUS: G9A deficient mice show a higher level of histone
CC       H3 with acetylated Lys-9 and/or methylated Lys-4, display severe
CC       developmental defects and die within E9.5-E12.5 stages.
CC   -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC       family. Suvar3-9 subfamily.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- SIMILARITY: Contains 1 pre-SET domain.
CC   -!- SIMILARITY: Contains 1 SET domain.
CC   -!- SIMILARITY: Contains 1 post-SET domain.
CC   -!- CAUTION: NG36 and G9a were originally thought to derive from two
CC       separate genes.
CC   -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC       gene model prediction.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF109906; AAC84164.1; ALT_SEQ.
DR   EMBL; AF109906; AAC84165.1; ALT_SEQ.
DR   EMBL; AB077209; BAC05482.1; -.
DR   EMBL; AB077210; BAC05483.1; -.
DR   EMBL; BC025539; AAH25539.1; ALT_INIT.
DR   MGD; MGI:2148922; Bat8.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR003616; PostSET.
DR   InterPro; IPR007728; Pre-SET.
DR   InterPro; IPR001214; SET.
DR   InterPro; IPR003606; Zn2-binding.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF05033; Pre-SET; 1.
DR   Pfam; PF00856; SET; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00468; PreSET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50868; POST_SET; FALSE_NEG.
DR   PROSITE; PS50867; PRE_SET; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW   ANK repeat; Repeat; Alternative splicing.
FT   DOMAIN      352    379       POLY-GLU.
FT   REPEAT      702    731       ANK 1.
```

```
FT   REPEAT      737    766       ANK 2.
FT   REPEAT      770    799       ANK 3.
FT   REPEAT      803    833       ANK 4.
FT   REPEAT      837    866       ANK 5.
FT   REPEAT      870    899       ANK 6.
FT   REPEAT      903    932       ANK 7.
FT   DOMAIN     1025   1088       PRE-SET.
FT   DOMAIN     1090   1212       SET.
FT   DOMAIN     1217   1233       POST-SET.
FT   VARSPLIC      1     57       Missing (in isoform 2).
FT                                /FTId=VSP_002214.
FT   VARSPLIC     58     71       AGLTGPPVPCLPSQ -> MAAAAGAAAAAAAE (in
FT                                isoform 2).
FT                                /FTId=VSP_002215.
FT   VARSPLIC    426    459       Missing (in isoform 2).
FT                                /FTId=VSP_002216.
FT   MUTAGEN    1162   1162       R->H: STRONGLY REDUCES HISTONE
FT                                METHYLTRANSFERASE ACTIVITY.
FT   MUTAGEN    1165   1168       MISSING: ABOLISHES HISTONE
FT                                METHYLTRANSFERASE ACTIVITY AND SUBSEQUENT
FT                                REPRESSION.
SQ   SEQUENCE   1263 AA;  138038 MW;  74DBFF9A36769589 CRC64;

  Query Match             4.5%;  Score 288;  DB 1;  Length 1263;
  Best Local Similarity   23.4%;  Pred. No. 4.8e-09;
  Matches  143;  Conservative   91;  Mismatches  229;  Indels  148;  Gaps   22;

Qy         83 GGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGF 142
              ||  | ||       |    | :|    : || |:|    : : |         |: :
Db        668 GG--CLSAVGLPPGPGREALEKALVIQESERRKKL----RFHPRQ--------LYLSVKQ 713

Qy        143 GR-KDVVEYLLQN-GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              |  : |:  || |   | |:       ||| |   |  |: ::||: ||: || |    |
Db        714 GELQKVILMLLDNLDPNFQSDQQSKRTPLHAAAQKGSVEICHVLLQAGANINAVDKQQRT 773

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              || || :   ::|   ::| |     :  || | |                        :|:
Db        774 PLMEAVVNNHLEVARYMVQLGGCVYSKEEDGSTCLH-------------------HAAK 813

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
               || | :  ||:   |: :| |    ||:  || :  : ::::||  ||||   |  : :
Db        814 IGNLEMVSLLLSTGQVDVNAQDSGGWTPIIWAAEHKHIDVIRMLLTRGADVTLTDNEENI 873

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              || |   |   : |:|:     ::|::    |||| || ::  :   | || ||:| |
Db        874 CLHWASFTGSAAIAEVLLNAQCDLHAVNYHGDTPLHIAARESYHDCVLLFLSRGANPELR 933

Qy        381 NCHNKSAIDLAP-------TPQLKERLAYEFKGHSLLQAARE--ADVTRIKKHLSLEMVN 431
              |    :| || |          ||  :|      |:  ::  :     || |  ::: :  ||
Db        934 NKEGDTAWDLTPERSDVWFALQLNRKLRLGV-GNRAVRTEKIICRDVARGYENVPIPCVN 992

Qy        432 ----------FKH-PQTHETAL-----------HCAAASPYPKRKQIC-ELLLR----KG 464
                       :|:  :  ||:              ||           :| :| :|     |
Db        993 GVDGEPCPEDYKYISENCETSTMNIDRNITHLQHCTCVDDCSSSNCLCGQLSIRCWYDKD 1052

Qy        465 ANINEKTKEFLTPLHVASEKA-------HNDVV--------------------------- 490
                : ::  :   ||      :|        | ||
Db       1053 GRLLQEFNKIEPPLIFECNQACSCWRSCKNRVVQSGIKVRLQLYRTAKMGWGVRALQTIP 1112

Qy        491 ----------EVVVKHEAKVNA-------LDNL-GQTSLHRAAYCGHLQTCRLLLSYGCD 532
                       |::   || |          |||  |:     | | |::      ::: ||
Db       1113 QGTFICEYVGELISDAEADVREDDSYLFDLDNKDGEVYCIDARYYGNISR---FINHLCD 1169

Qy        533 PNIISLQGFTALQMGNENVQ---------QLLQEGISLGNSEADRQLLEAAKAGDVETVK 583
              |||| :: |    | ::::          : :: |  ||    ||          |:::
Db       1170 PNIIPVRVF----MLHQDLRFPRIAFFSSRDIRTGEELGFDYGDR-------FWDIKSKY 1218

Qy        584 KLCTVQSVNCR 594
                |   |  |:
Db       1219 FTCQCGSEKCK 1229

Search completed: March 25, 2004, 16:42:41
```

Job time : 37.8376 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:39:11 ; Search time 22.2564 Seconds
(without alignments)
2876.304 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6464	100.0	1240	4	US-09-696-668-4	Sequence 4, Appli
2	6108	94.5	1166	4	US-09-972-115A-6	Sequence 6, Appli
3	6053	93.6	1166	4	US-09-350-982C-5	Sequence 5, Appli
4	5766	89.2	1100	4	US-09-696-668-3	Sequence 3, Appli
5	5731.5	88.7	1100	4	US-09-427-154-2	Sequence 2, Appli
6	5452	84.3	1267	4	US-09-972-115A-4	Sequence 4, Appli
7	5103.5	79.0	1327	3	US-09-196-387-2	Sequence 2, Appli
8	5103.5	79.0	1327	4	US-09-841-835-2	Sequence 2, Appli
9	5103.5	79.0	1327	4	US-09-972-115A-8	Sequence 8, Appli
10	5031.5	77.8	1333	4	US-09-972-115A-2	Sequence 2, Appli
11	3307	51.2	949	3	US-09-196-387-10	Sequence 10, Appl
12	3307	51.2	949	4	US-09-841-835-10	Sequence 10, Appl
13	2132.5	33.0	673	3	US-09-196-387-8	Sequence 8, Appli
14	2132.5	33.0	673	4	US-09-841-835-8	Sequence 8, Appli
15	877	13.6	1839	2	US-09-172-977-4	Sequence 4, Appli
16	877	13.6	1839	4	US-09-404-108-4	Sequence 4, Appli
17	874	13.5	160	4	US-09-972-115A-20	Sequence 20, Appl
18	858	13.3	843	2	US-09-172-977-3	Sequence 3, Appli
19	858	13.3	843	4	US-09-404-108-3	Sequence 3, Appli
20	835	12.9	1745	2	US-09-031-485-33	Sequence 33, Appl
21	835	12.9	1745	2	US-08-847-429A-33	Sequence 33, Appl
22	835	12.9	1745	3	US-09-065-474-33	Sequence 33, Appl
23	835	12.9	1745	4	US-09-557-034-33	Sequence 33, Appl
24	732	11.3	994	4	US-10-164-595-38	Sequence 38, Appl
25	633.5	9.8	1619	4	US-09-392-812A-4	Sequence 4, Appli
26	629.5	9.7	1088	3	US-09-082-059-2	Sequence 2, Appli
27	601.5	9.3	683	4	US-10-164-595-71	Sequence 71, Appl
28	563	8.7	1704	4	US-09-392-812A-2	Sequence 2, Appli
29	545.5	8.4	1709	4	US-09-392-812A-6	Sequence 6, Appli
30	514.5	8.0	352	3	US-09-065-474-139	Sequence 139, App
31	514.5	8.0	352	4	US-09-557-034-139	Sequence 139, App
32	440	6.8	90	4	US-09-972-115A-22	Sequence 22, Appl
33	421.5	6.5	303	2	US-09-031-485-23	Sequence 23, Appl
34	421.5	6.5	303	2	US-08-847-429A-23	Sequence 23, Appl
35	421.5	6.5	303	3	US-09-065-474-23	Sequence 23, Appl
36	421.5	6.5	303	4	US-09-557-034-23	Sequence 23, Appl
37	409	6.3	302	2	US-09-031-485-38	Sequence 38, Appl
38	409	6.3	302	2	US-08-847-429A-38	Sequence 38, Appl
39	409	6.3	302	3	US-09-065-474-38	Sequence 38, Appl
40	409	6.3	302	4	US-09-557-034-38	Sequence 38, Appl
41	398.5	6.2	1719	4	US-09-439-313-378	Sequence 378, App
42	398.5	6.2	1719	4	US-09-352-616A-378	Sequence 378, App
43	398.5	6.2	1719	4	US-09-636-215-378	Sequence 378, App
44	398.5	6.2	1719	4	US-09-685-166A-378	Sequence 378, App
45	392	6.1	348	2	US-09-031-485-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-696-668-4
; Sequence 4, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Chan, Eva
; APPLICANT: Xu, Xiang
; APPLICANT: Huang, Betty
; APPLICANT: Ossovskaya, Valeria
; TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS (
; FILE REFERENCE: A-68292-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/696,668
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/427,154
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-696-668-4

Query Match 100.0%; Score 6464; DB 4; Length 1240;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240

Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420

Qy        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080

Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||
Db       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
```

```
RESULT 2
US-09-972-115A-6
; Sequence 6, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 1166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-6

  Query Match             94.5%;  Score 6108;  DB 4;  Length 1166;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1166;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy         135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy         195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy         255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy         315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy         375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy         435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy         495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy         555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy         615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy         675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy         735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy         795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy         855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166


RESULT 3
US-09-350-982C-5
; Sequence 5, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
;  APPLICANT:  Berthelsen, Jens
;  APPLICANT:  Toma, Salvatore
;  APPLICANT:  Isacchi, Antonella
;  TITLE OF INVENTION:  Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Rela
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  PHRM-0043
;  CURRENT APPLICATION NUMBER: US/09/350,982C
;  CURRENT FILING DATE:  1999-07-09
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 5
;  LENGTH: 1166
;  TYPE: PRT
;  ORGANISM: Artificial
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Xaa is any amino acid
;  NAME/KEY: misc_feature
;  LOCATION: (1102)..(1102)
;  OTHER INFORMATION: n is any nucleic acid
;  NAME/KEY: misc_feature
;  LOCATION: (2650)..(2650)
;  OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-5

  Query Match           93.6%;  Score 6053;  DB 4;  Length 1166;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1157;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy         135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PLHLAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy         195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy         255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy         315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy         375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy         435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
               ||||||| |||||||||||||||||||||| |||||||||||||||||||||| |||||
Db         361 PQTHETAXHCAAASPYPKRKQICELLLRKGAXINEKTKEFLTPLHVASEKAHNDXVEVVV 420

Qy         495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy         555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy         615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
               |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db         541 VVEYLLQHGADVHAKDKXXLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy         675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKSCLAR 660

Qy         735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy         795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
               |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALXLAHGADPTLKNQEGQT 780

Qy         855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy         915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
               |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRXLGLEHLMDIFEREQIT 900

Qy         975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy        1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy        1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy        1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy        1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
               |||||||||||||||||||||||||
Db        1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166


RESULT 4
US-09-696-668-3
; Sequence 3, Application US/09696668
; Patent No. 6617102
```

```
; GENERAL INFORMATION:
;  APPLICANT: Luo, Ying
;  APPLICANT:  Chan, Eva
;  APPLICANT:  Xu, Xiang
;  APPLICANT:  Huang, Betty
;  APPLICANT:  Ossovskaya, Valeria
;  TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS
;  FILE REFERENCE: A-68292-1/RMS/DHR
;  CURRENT APPLICATION NUMBER: US/09/696,668
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 09/427,154
;  PRIOR FILING DATE: 1999-10-25
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1100
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-696-668-3

  Query Match             89.2%;  Score 5766;  DB 4;  Length 1100;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1221 AYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100


RESULT 5
US-09-427-154-2
; Sequence 2, Application US/09427154
; Patent No. 6589725
; GENERAL INFORMATION:
;  APPLICANT:  Luo, Ying
;  APPLICANT:  Chan, Eva
;  APPLICANT:  Xu, Xiang
;  APPLICANT:  Huang, Betty
;  TITLE OF INVENTION:  TANKYRASE H, COMPOSTIONS INVOLVED IN THE CELL CYCLE AND
;  TITLE OF INVENTION:  METHODS OF USE
;  FILE REFERENCE:  A-68292-DJB/RMS/DAV
;  CURRENT APPLICATION NUMBER: US/09/427,154
;  CURRENT FILING DATE:  1999-10-25
;  NUMBER OF SEQ ID NOS:  9
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1100
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-427-154-2

  Query Match             88.7%;  Score 5731.5;  DB 4;  Length 1100;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1095;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTILNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
```

```
Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIHYLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSLVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 -SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1079

Qy       1221 AYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||
Db       1080 AYPEYLITYQIMRPEGMVDG 1099
```

RESULT 6
US-09-972-115A-4
; Sequence 4, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: The 'Xaa' at location 1 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: The 'Xaa' at location 42 stands for Lys, Arg, Thr, or Met.
; NAME/KEY: misc_feature
; LOCATION: (98)..(98)
; OTHER INFORMATION: The 'Xaa' at location 98 stands for Lys, Arg, Thr, or Met.
; NAME/KEY: misc_feature
; LOCATION: (110)..(110)
; OTHER INFORMATION: The 'Xaa' at location 110 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (114)..(114)
; OTHER INFORMATION: The 'Xaa' at location 114 stands for Asn, Ser, Thr, or Ile.
; NAME/KEY: misc_feature
; LOCATION: (126)..(126)
; OTHER INFORMATION: The 'Xaa' at location 126 stands for Lys, or Asn.
; NAME/KEY: misc_feature
; LOCATION: (139)..(139)
; OTHER INFORMATION: The 'Xaa' at location 139 stands for Ile, Val, Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (144)..(144)
; OTHER INFORMATION: The 'Xaa' at location 144 stands for Thr.
; NAME/KEY: misc_feature
; LOCATION: (153)..(153)
; OTHER INFORMATION: The 'Xaa' at location 153 stands for Thr.
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: The 'Xaa' at location 166 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: The 'Xaa' at location 173 stands for Ile, Val, Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (174)..(174)
; OTHER INFORMATION: The 'Xaa' at location 174 stands for Glu, or Lys.
; NAME/KEY: misc_feature
; LOCATION: (183)..(183)
; OTHER INFORMATION: The 'Xaa' at location 183 stands for Val.
; NAME/KEY: misc_feature
; LOCATION: (185)..(185)
; OTHER INFORMATION: The 'Xaa' at location 185 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: The 'Xaa' at location 188 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (202)..(202)
; OTHER INFORMATION: The 'Xaa' at location 202 stands for His, or Pro.
; NAME/KEY: misc_feature
; LOCATION: (204)..(204)
; OTHER INFORMATION: The 'Xaa' at location 204 stands for Glu, or Lys.
; NAME/KEY: misc_feature
; LOCATION: (205)..(205)
; OTHER INFORMATION: The 'Xaa' at location 205 stands for Asn, Ser, Thr, or Ile.
; NAME/KEY: misc_feature
; LOCATION: (207)..(207)
; OTHER INFORMATION: The 'Xaa' at location 207 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (211)..(211)
; OTHER INFORMATION: The 'Xaa' at location 211 stands for Gln, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (213)..(213)
; OTHER INFORMATION: The 'Xaa' at location 213 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature

; LOCATION: (215)..(215)
; OTHER INFORMATION: The 'Xaa' at location 215 stands for Lys, Asn, Glu, Asp, Gln, His
; OTHER INFORMATION: , a stop codon, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (227)..(227)
; OTHER INFORMATION: The 'Xaa' at location 227 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (228)..(228)
; OTHER INFORMATION: The 'Xaa' at location 228 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (237)..(237)
; OTHER INFORMATION: The 'Xaa' at location 237 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (238)..(238)
; OTHER INFORMATION: The 'Xaa' at location 238 stands for Asn, Ser, Thr, Ile, Asp, Gly
; OTHER INFORMATION: , Ala, Val, His, Arg, Pro, Leu, Tyr, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (667)..(667)
; OTHER INFORMATION: The 'Xaa' at location 667 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (725)..(725)
; OTHER INFORMATION: The 'Xaa' at location 725 stands for Lys, or Asn.
; NAME/KEY: misc_feature
; LOCATION: (727)..(727)
; OTHER INFORMATION: The 'Xaa' at location 727 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (735)..(735)
; OTHER INFORMATION: The 'Xaa' at location 735 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (741)..(741)
; OTHER INFORMATION: The 'Xaa' at location 741 stands for Met, Val, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (763)..(763)
; OTHER INFORMATION: The 'Xaa' at location 763 stands for Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (764)..(764)
; OTHER INFORMATION: The 'Xaa' at location 764 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (787)..(787)
; OTHER INFORMATION: The 'Xaa' at location 787 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (788)..(788)
; OTHER INFORMATION: The 'Xaa' at location 788 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (789)..(789)
; OTHER INFORMATION: The 'Xaa' at location 789 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (790)..(790)
; OTHER INFORMATION: The 'Xaa' at location 790 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (791)..(791)
; OTHER INFORMATION: The 'Xaa' at location 791 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (792)..(792)
; OTHER INFORMATION: The 'Xaa' at location 792 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (793)..(793)
; OTHER INFORMATION: The 'Xaa' at location 793 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (794)..(794)
; OTHER INFORMATION: The 'Xaa' at location 794 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (795)..(795)
; OTHER INFORMATION: The 'Xaa' at location 795 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (796)..(796)
; OTHER INFORMATION: The 'Xaa' at location 796 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: The 'Xaa' at location 797 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (798)..(798)
; OTHER INFORMATION: The 'Xaa' at location 798 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (799)..(799)
; OTHER INFORMATION: The 'Xaa' at location 799 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (800)..(800)
; OTHER INFORMATION: The 'Xaa' at location 800 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (801)..(801)
; OTHER INFORMATION: The 'Xaa' at location 801 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (802)..(802)

Query Match 84.3%; Score 5452; DB 4; Length 1267;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 1072; Conservative 6; Mismatches 161; Indels 4; Gaps 2;

```
Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
              ||  |||||||||| |    | |||||||||||||||||||||||||||||||||||||
Db         26 RCLRRRGAAGGQGAHRXGARGRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 85

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
              |||||||||||| |||||||||| ||| |||||||||| |||||||||||| :|| |
Db         86 RLLALLLAVAAAXIMSGRRCAGGGXACAXAAAEAVEPAARXLFEACRNGDVERXKKLVXP 145

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
              ||||||| |||||||||| | |||||:  ||| |||| | || ||||||||||||| |
Db        146 EKVNSRDXAGRKSTPLHFPAXFGRKDLXXYLLTNGANXQXRDXGGLIPLHNACSFGXAXX 205

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
              : ||| | | |||||||||||  ||||||||  ||||||||||||||||||||||||||
Db        206 IXLLLXHXAXPNARDNWNYTPXXEAAIKGKIXXCIVLLQHGAEPTIRNTDGRTALDLADP 265

Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 325

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 385

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 445
```

```
Qy        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 505

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 565

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 625

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        626 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVN 685

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
              |||||||||||||||||||||||||||||||| :  |  ||||||| ||||| ||||
Db        686 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKKKKGXIXLDLVKDGXTDIQDXLRGD 745

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
              | ||||||||||||||   ||||||||||||||||||| ||
Db        746 AVLLDAAKKGCLARVKKXXFPDNVNCRDTQGRHSTPLHL-AGXXXXXXXXXXXXXXXXXX 804

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840

Db        805 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 864

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
                                          ||||||| |||| ||||||||||||||||||
Db        865 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALS 924

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 984

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
              |||||||| |||||||||||||||||||| |||||||||||  |||||||||||||||||
Db        985 LEHLMDIFXREQITLDVLVEMGHKELKEIXINAYGHRHKLIKSFERLISGQQGLNPYLTL 1044

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK--- 1077
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db       1045 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNRAKI 1104

Qy       1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
              :  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 RHEERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1164

Qy       1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1224

Qy       1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||||||||||||||||||||
Db       1225 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1267


RESULT 7
US-09-196-387-2
; Sequence 2, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-196-387-2

  Query Match           79.0%;  Score 5103.5;  DB 3;  Length 1327;
  Best Local Similarity  79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels  15;  Gaps  3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||     : |   |  :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||        ||  | ||| |||||||| |||||    ||::| |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||||:|||  ||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| |||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||:|||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||||:|||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:|||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||||| ||:|||||||| ||||| |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
```

```
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681

Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741

Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801

Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861

Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921

Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979

Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039

Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099

Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159

Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219

Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235

Db       1299 QAYPEYLITYQIMKPE 1314


RESULT 8
US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-841-835-2

  Query Match           79.0%;  Score 5103.5;  DB 4;  Length 1327;
  Best Local Similarity  79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81

Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141

Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201

Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261

Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321

Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381

Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441

Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501

Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561

Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621

Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681

Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741

Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825
```

```
Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235
Db       1299 QAYPEYLITYQIMKPE 1314


RESULT 9
US-09-972-115A-8
; Sequence 8, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 1327
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-8

  Query Match           79.0%;  Score 5103.5;  DB 4;  Length 1327;
  Best Local Similarity   79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298
```

```
Qy       1220 QAYPEYLITYQIMRPE 1235
              ||||||||||||||:||
Db       1299 QAYPEYLITYQIMKPE 1314


RESULT 10
US-09-972-115A-2
; Sequence 2, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1333
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: The 'Xaa' at location 1 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: The 'Xaa' at location 2 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: The 'Xaa' at location 3 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: The 'Xaa' at location 4 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: The 'Xaa' at location 5 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (6)..(6)
;   OTHER INFORMATION: The 'Xaa' at location 6 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: The 'Xaa' at location 7 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: The 'Xaa' at location 8 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: The 'Xaa' at location 9 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: The 'Xaa' at location 10 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (11)..(11)
;   OTHER INFORMATION: The 'Xaa' at location 11 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (12)..(12)
;   OTHER INFORMATION: The 'Xaa' at location 12 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: The 'Xaa' at location 13 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (14)..(14)
;   OTHER INFORMATION: The 'Xaa' at location 14 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (15)..(15)
;   OTHER INFORMATION: The 'Xaa' at location 15 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (16)..(16)
;   OTHER INFORMATION: The 'Xaa' at location 16 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (17)..(17)
;   OTHER INFORMATION: The 'Xaa' at location 17 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (19)..(19)
;   OTHER INFORMATION: The 'Xaa' at location 19 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (20)..(20)
;   OTHER INFORMATION: The 'Xaa' at location 20 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (21)..(21)
;   OTHER INFORMATION: The 'Xaa' at location 21 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (22)..(22)
;   OTHER INFORMATION: The 'Xaa' at location 22 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (23)..(23)
;   OTHER INFORMATION: The 'Xaa' at location 23 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: The 'Xaa' at location 24 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (25)..(25)
;   OTHER INFORMATION: The 'Xaa' at location 25 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (26)..(26)
;   OTHER INFORMATION: The 'Xaa' at location 26 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (27)..(27)
;   OTHER INFORMATION: The 'Xaa' at location 27 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (28)..(28)
;   OTHER INFORMATION: The 'Xaa' at location 28 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(29)
;   OTHER INFORMATION: The 'Xaa' at location 29 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: The 'Xaa' at location 30 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: The 'Xaa' at location 31 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(32)
;   OTHER INFORMATION: The 'Xaa' at location 32 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (33)..(33)
;   OTHER INFORMATION: The 'Xaa' at location 33 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (34)..(34)
;   OTHER INFORMATION: The 'Xaa' at location 34 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (35)..(35)
;   OTHER INFORMATION: The 'Xaa' at location 35 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: The 'Xaa' at location 36 stands for Lys, Asn, Arg, Ser, Thr, Ile,

  Query Match             77.8%;  Score 5031.5;  DB 4;  Length 1333;
  Best Local Similarity   86.7%;  Pred. No. 0;
  Matches  983;  Conservative    7;  Mismatches  135;  Indels    9;  Gaps    2;

Qy        116 RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSF 175
              |||||||||  :| |||||||||  ||||||:  :||||||| |   :|||||| ||||
Db        200 RLVTPEKVNXXNTXGRKSTPLHFPAGFGRKNXXKYLLQNGANXQXLYNGGLIPLHXACSF 259

Qy        176 GHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTAL 235
              |||: :|||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        260 GHAKXINLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTAL 319

Qy        236 DLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY 295
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 DLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY 379

Qy        296 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 355
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 439

Qy        356 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 415
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 499

Qy        416 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 475
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 559

Qy        476 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 535
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 619

Qy        536 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 595
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 679

Qy        596 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 655
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKH 739

Qy        656 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD 715
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTLLDLVKDGXTDIQD 799

Qy        716 LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA--AGYNNLEVAEYLL 773
              ||||| |||||||||||||||   ||||||||||||||||||||
Db        800 XLRGDAVLLDAAKKGCLARVKKXXFPDNVNCRDTQGRHSTPLHLAGXXXXXXXXXXXXXXX 859

Qy        774 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 919

Qy        834 LCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPG 893
              |||||||| |||| ||||||||||||
Db        920 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPG 979

Qy        894 ATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSIT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 ATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSIT 1039

Qy        954 QFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 1013
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db       1040 QFVRNLGLEHLMDIFXREQITLDVLVEMGHKELKEIXINAYGHRHKLIKGVERLISGQQG 1099

Qy       1014 LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQK 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQK 1159

Qy       1074 VCNK-------KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1126
              |||         :|||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 VCNXXXXXAKIRHEERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1219

Qy       1127 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1279

Qy       1187 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1333

RESULT 11
US-09-196-387-10
```

```
; Sequence 10, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 949 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-196-387-10

  Query Match             51.2%;  Score 3307;  DB 3;  Length 949;
  Best Local Similarity   76.2%;  Pred. No. 2.5e-279;
  Matches  638;  Conservative   72;  Mismatches  119;  Indels    8;  Gaps    2;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |     | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||         ||  | ||| |||||||| ||||||    ||:: | |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              |||||||:||| |||| ||||||||||||||||||||:|||  ||||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||

Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||||| ||:|||||||| ||||| |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
                |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              |||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||:|||||||:|||||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
              :|:||||||||:|||||||||||||||||||:|||||||||||||||||||||||
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCLA 885

Qy        800 DVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856
              |  |    |    |    |||| ||      ::   || ||||   ::: |  ||
Db        886 RVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942


RESULT 12
US-09-841-835-10
; Sequence 10, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 949 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
US-09-841-835-10

  Query Match             51.2%;  Score 3307;  DB 4;  Length 949;
  Best Local Similarity   76.2%;  Pred. No. 2.5e-279;
  Matches  638;  Conservative   72;  Mismatches  119;  Indels    8;  Gaps    2;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :|| : : |:|     :   : ||      : |     |  :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||        ||  | ||| |||||||| ||||||    ||::| |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| ||||||||||||||||||||:||| |||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| | :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
                |: | :||||:||||:||||:||: |:|||||:||| |||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              |||||||||||||||||||||||||||||:||| |||||||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||||:|||||||||:|||||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
              :|:|||||||:|||||||||||||||||:|||||||||||||||||||||||||
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCLA 885

Qy        800 DVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856
              |  |   |    |     |||| ||     ::   || ||||   ::: |  ||
Db        886 RVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942


RESULT 13
US-09-196-387-8
; Sequence 8, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 673 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-196-387-8

  Query Match             33.0%;  Score 2132.5;  DB 3;  Length 673;
  Best Local Similarity   73.3%;  Pred. No. 4.9e-177;
  Matches  417;  Conservative   54;  Mismatches   91;  Indels    7;  Gaps    2;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :|| : : |:|     :   : ||      : |     |  :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||        ||  | ||| |||||||| ||||||    ||::| |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| ||||||||||||||||||||:||| |||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
```

```
                |||  ||||:|||||  ||||||||||||| ||:|||||||| |||||  |||:|
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYGS 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQS 590
                |    |   ||: |: | |::: :  |
Db        646 DPSIISLQGFTAAQMGN-EAVQQILSGHS 673


RESULT 14
US-09-841-835-8
; Sequence 8, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 673 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-841-835-8

  Query Match             33.0%;  Score 2132.5;  DB 4;  Length 673;
  Best Local Similarity   73.3%;  Pred. No. 4.9e-177;
  Matches  417;  Conservative   54;  Mismatches   91;  Indels    7;  Gaps    2;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |      | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||         ||  | ||| |||||||| ||||||    ||::| |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| ||||||||||||||||||||:||| ||||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| |||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||||||:|||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              |||  ||||:|||||  |||||||||||||| ||:||||||||| ||||| |||:|
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYGS 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQS 590
                |    |   ||: |: | |::: :  |
Db        646 DPSIISLQGFTAAQMGN-EAVQQILSGHS 673


RESULT 15
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Yue, Henry
;  TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
;  FILE REFERENCE: PF-0615 US
;  CURRENT APPLICATION NUMBER: US/09/172,977
;  CURRENT FILING DATE: 1998-10-14
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PERL Program
; SEQ ID NO 4
;   LENGTH: 1839
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: g29491
US-09-172-977-4

  Query Match             13.6%;  Score 877;  DB 2;  Length 1839;
  Best Local Similarity   29.7%;  Pred. No. 1.2e-66;
  Matches  293;  Conservative  137;  Mismatches  402;  Indels  156;  Gaps   23;

Qy        105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
              | | |:::: |: |     :|: :  |  :  || ||  |   :|: ||  |::| :
Db         38 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 95

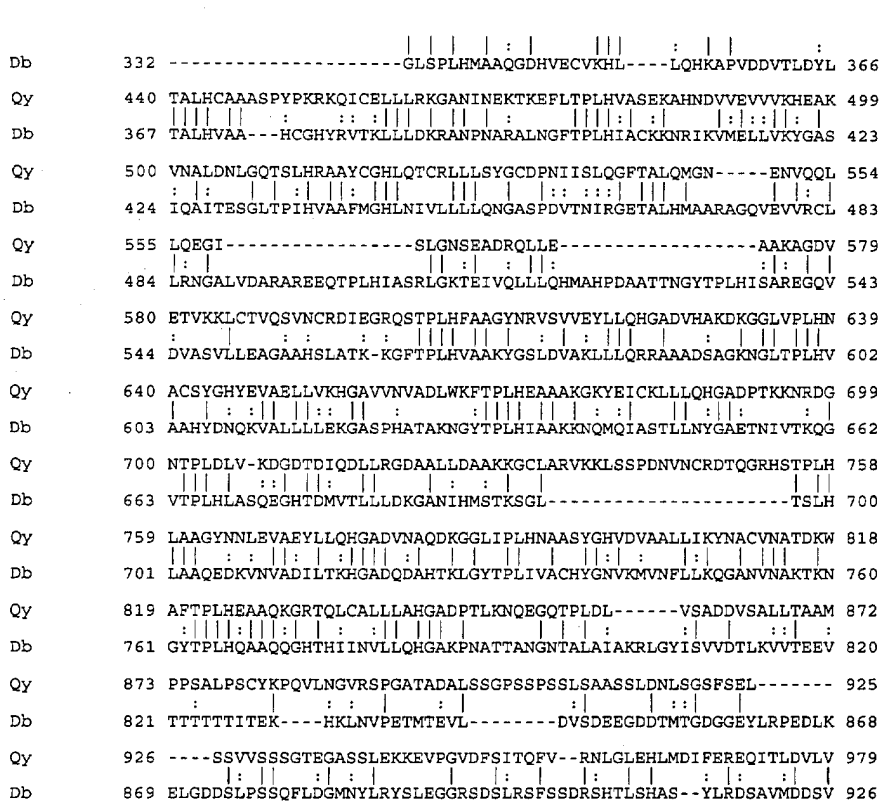
Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
              |   || |   | |||| :|:: ||: ||:    :|||: || :  |||   ||::||
Db         96 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 155

Qy        224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
              .:    || |    ||      | |:|     | ::    |  |   :: |  |||
Db        156 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 213

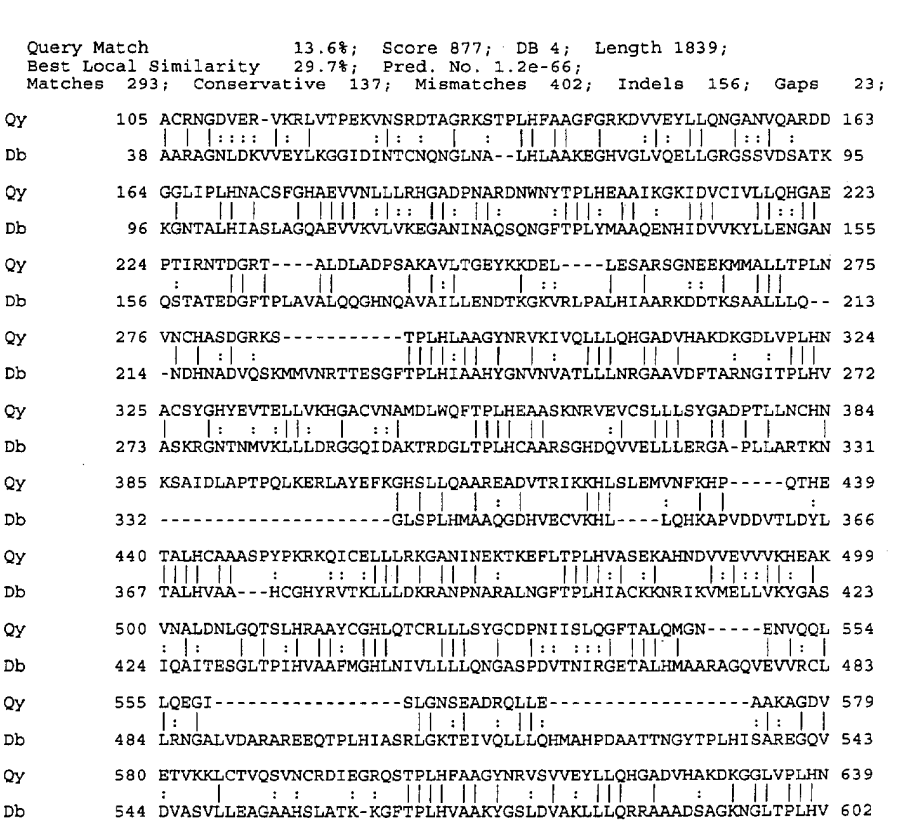
Qy        276 VNCHASDGRKS----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
              | | :| :            ||||:|| |  | :  |||  || |    :  : |||
Db        214 -NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 272

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
              |   |:  : :||:  |  ::|     ||||  ||     :|  |||  || | |     |
Db        273 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 331

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
```



RESULT 16
US-09-404-108-4
; Sequence 4, Application US/09404108
; Patent No. 6590077
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/404,108
; CURRENT FILING DATE: 1999-09-23
; EARLIER APPLICATION NUMBER: 09/172,977
; EARLIER FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
US-09-404-108-4



RESULT 17
US-09-972-115A-20
; Sequence 20, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-20

Query Match 13.5%; Score 874; DB 4; Length 160;
Best Local Similarity 100.0%; Pred. No. 3.1e-68;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 60

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 120

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
             ||||||||||||||||||||||||||||||||||||||||
Db        121 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 160
```

RESULT 18
US-09-172-977-3
; Sequence 3, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: g1841966
US-09-172-977-3

Query Match 13.3%; Score 858; DB 2; Length 843;
Best Local Similarity 29.6%; Pred. No. 1.4e-65;
Matches 275; Conservative 148; Mismatches 367; Indels 140; Gaps 23;

```
Qy        105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
              | | |:::: |: |      :|: :  |  :  || ||  |   :|: ||  |::| :
Db          3 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 60

Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
              |   || |   | |||| :|:: ||: ||:    :|||: || :  |||    ||::||
Db         61 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 120

Qy        224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
              :  || |   ||     | |:|    | ::    |  |   :: |  |||
Db        121 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 178

Qy        276 VNCHASDGRKS----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
               | | :| :           ||||:|| |  | : |||  || |   :  : |||
Db        179 -NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 237

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
              |   |:  : :||:  |  ::|      |||| ||     :|  |||  || | |    |
Db        238 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 296

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
                                  | | |  | : |     |||    :  | |       :
Db        297 ---------------------GLSPLHMAAQGDHVECVKHL----LQHKAPVDDVTLDYL 331

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              |||| ||   :    :: :||| | || | :     ||||:| :|    |:|::||: |
Db        332 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAY 388

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |:   | | :  ||: |||    |||  |  |:: :::| ||| |
Db        389 IQAITESGLTPIPVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHM------------ 435

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
                          ||:||:|| |:  | :::    |   | : |||| |:   :  :|:
Db        436 ------------AARAGEVEVVR--CLLRNGALVDARAREEQTPLHIASRLGKTEIVQL 480

Qy        619 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 678
              |||| |   |    |  ||| :   |  :|| :|:: ||  ::|    |||| ||  |
Db        481 LLQHMAHPDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATKKGFTPLHVAAKYGS 540

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              :: |||||  |      ::| ||| :    |                  :|  |  ::|
Db        541 LDVAKLLLQRRAAADSAGKNGLTPLHVAAHYDN----------------QKVALLLLEKG 584

Qy        739 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 798
              :||        |    ||||:||  | :::|  || :||: |   | |: ||| |:  ||
Db        585 ASP-----HATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNTVTKQGVTPLHLASQEGH 639

Qy        799 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 858
              |:  |:::  | :: : |   | || ||:: :  :  :|  ||||     : | ||| :
Db        640 TDMVTLVLEKGANIHMSTKSGLTSLHLAAEEDKVNVADILTKHGADQDAYTKLGYTPLIV 699

Qy        859 V----SADDVSALLTAAMPPSA-LPSCYKP----------QVLNGVRSPGATADALSSGP 903
                  :  |: ||      :|  :  | |           ::| :   ||  :| ::
Db        700 ACHYGNVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTA-- 757

Qy        904 SSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEH 963
              :  ::|: |  |  :  |  :   ||:    |  ::: :|          ::|:
Db        758 NGNTALAIAKRLGYI--SVVDTLKVVTEEVTTTTTTITEKHKLNAPETMTE--------- 806

Qy        964 LMDIFERE---QITLDVLVEMGHKELKEIG 990
              ::|: : |    :| |    :  ::|||:|
Db        807 VLDVSDEEGDDTVTGDGGEYLRPEDLKELG 836
```

RESULT 19
US-09-404-108-3
; Sequence 3, Application US/09404108
; Patent No. 6590077
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/404,108

```
;  CURRENT FILING DATE: 1999-09-23
;  EARLIER APPLICATION NUMBER: 09/172,977
;  EARLIER FILING DATE: 1998-10-14
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PERL Program
; SEQ ID NO 3
;   LENGTH: 843
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
;   FEATURE: -
;   OTHER INFORMATION: g1841966
US-09-404-108-3

  Query Match             13.3%;  Score 858;  DB 4;  Length 843;
  Best Local Similarity   29.6%;  Pred. No. 1.4e-65;
  Matches  275;  Conservative  148;  Mismatches  367;  Indels  140;  Gaps   23;

Qy        105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
              | |.|:::: |: |     :|: :  |  :  || ||  |   :|: ||  |::| :
Db          3 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 60

Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
              |   || |   | |||| :|:: ||: ||:    :|||: || :  |||   ||::||
Db         61 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 120

Qy        224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
              :    || |    ||       | |:|      | ::    |  |   :: |  |||
Db        121 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 178

Qy        276 VNCHASDGRKS----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
              | | :| :           ||||:|| |  | :  |||  || |    :  : |||
Db        179 -NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 237

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
              |   |:  : :||:  |  ::|      |||| ||     :|  |||  || | |    |
Db        238 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 296

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
                          | | |  | : |     |||    :  | |       :
Db        297 ----------------------GLSPLHMAAQGDHVECVKHL----LQHKAPVDDVTLDYL 331

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              |||| ||   :    :: :|| | || |  :      ||||:| :|    |:|::||: |
Db        332 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAY 388

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |:   | | :  ||: |||     |||  |  |:: :::| ||| |
Db        389 IQAITESGLTPIPVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHM------------ 435

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
                        ||:||:|| |:  | :::    |   | : |||| |:   :  :|:
Db        436 ------------AARAGEVEVVR--CLLRNGALVDARAREEQTPLHIASRLGKTEIVQL 480

Qy        619 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 678
              |||| |   |    |  ||| :   |  :|| :|:: ||  ::|     ||||| ||  |
Db        481 LLQHMAHPDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATKKGFTPLHVAAKYGS 540

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              :: |||||  |       ::| ||| :    |                    :|  |  ::|
Db        541 LDVAKLLLQRRAAADSAGKNGLTPLHVAAHYDN----------------QKVALLLLEKG 584

Qy        739 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 798
              :||       |    ||||:|| | :::|  || :||: |   | |: ||| |:  ||
Db        585 ASP-----HATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNTVTKQGVTPLHLASQEGH 639

Qy        799 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 858
              |:  |:::  | :: : |   | || ||:: :  :  :|  ||||    : | |||  :
Db        640 TDMVTLVLEKGANIHMSTKSGLTSLHLAAEEDKVNVADILTKHGADQDAYTKLGYTPLIV 699

Qy        859 V----SADDVSALLTAAMPPSA-LPSCYKP----------QVLNGVRSPGATADALSSGP 903
                    :   |: ||      :|   : | |           ::| :   ||  :| ::
Db        700 ACHYGNVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTA-- 757

Qy        904 SSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEH 963
              :  ::|: |  |  :  |  :   ||:   |   ::: :|       ::|:
Db        758 NGNTALAIAKRLGYI--SVVDTLKVVTEEVTTTTTTITEKHKLNAPETMTE--------- 806

Qy        964 LMDIFERE---QITLDVLVEMGHKELKEIG 990
              ::|: : |    :| |    :  ::|||:|
Db        807 VLDVSDEEGDDTVTGDGGEYLRPEDLKELG 836


RESULT 20
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1745 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-031-485-33

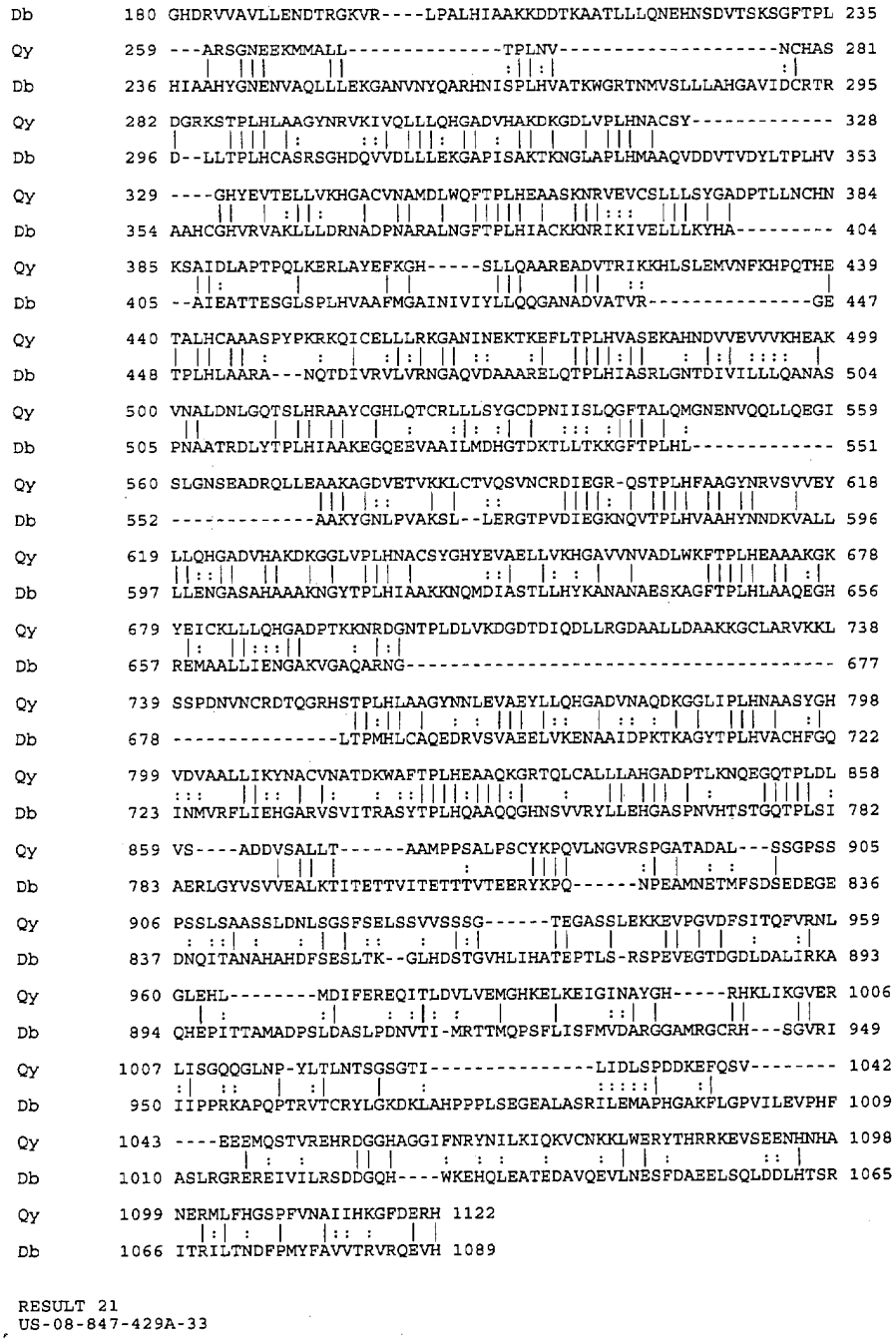
  Query Match             12.9%;  Score 835;  DB 2;  Length 1745;
  Best Local Similarity   25.7%;  Pred. No. 5e-63;
  Matches  314;  Conservative  154;  Mismatches  452;  Indels  304;  Gaps   37;

Qy         34 GSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSG---RRC-AGGGAACAS 89
              |  :|  | |:  | |           |:| ||          ||     | | |  |
Db         35 GESSASFLRAARAGNLD----------RVLELL--------RSGTDINTCNANGLNALHL 76

Qy         90 AAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRK-STPLHFAAGFGRKDVV 148
              |: |       ||| :  |  ||                 | | || :| || |:  |:: :|
Db         77 ASKEGHHEVVRELLK--RKADV--------------DAATRKGNTALHIASLAGQELIV 119

Qy        149 EYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIK 208
                |::|||||  :   |  ||: |     |  ||  || | |:        :|||  |  :
Db        120 TVLVENGANVNVQSLNGFTPLYMAAQENHESVVRYLLAHNANQALSTEDGFTPLAVALQQ 179

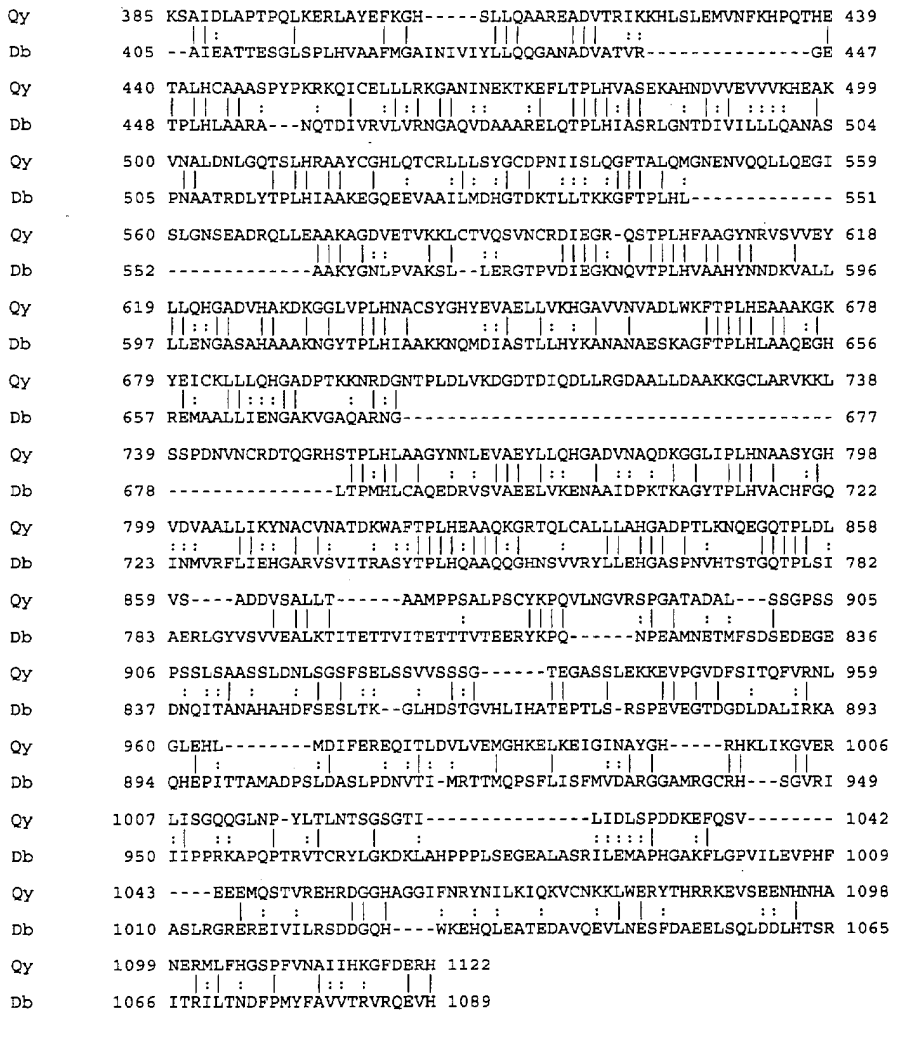
Qy        209 GKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYKKDELLES----- 258
              |   |  |||::     :|      || :|       :|  :|  |:  |   :|
```



RESULT 21
US-08-847-429A-33

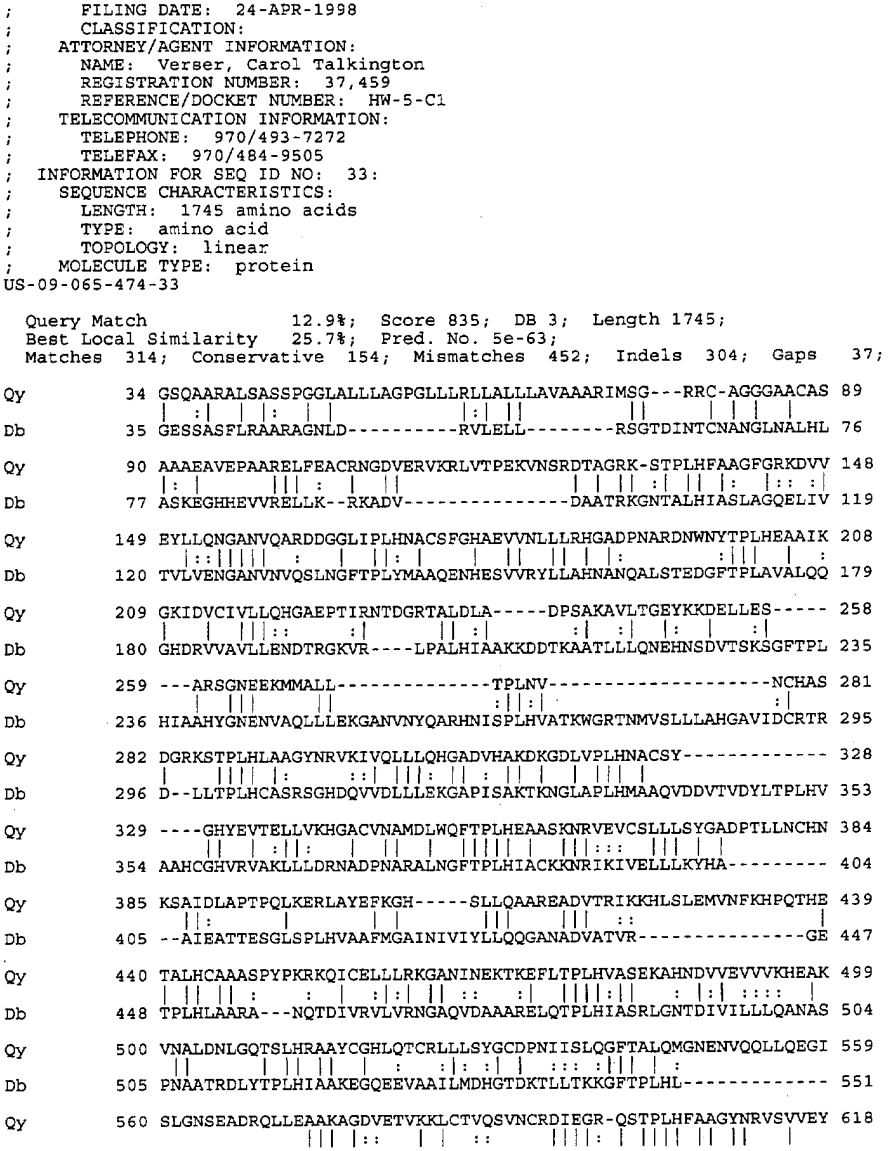
; Sequence 33, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-847-429A-33

Query Match 12.9%; Score 835; DB 2; Length 1745;
Best Local Similarity 25.7%; Pred. No. 5e-63;
Matches 314; Conservative 154; Mismatches 452; Indels 304; Gaps 37;



RESULT 22
US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-33

Query Match 12.9%; Score 835; DB 3; Length 1745;
Best Local Similarity 25.7%; Pred. No. 5e-63;
Matches 314; Conservative 154; Mismatches 452; Indels 304; Gaps 37;



```
Db        552 -------------AAKYGNLPVAKSL--LERGTPVDIEGKNQVTPLHVAAHYNNDKVALL 596

Qy        619 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 678
              ||::||  ||  | |  ||| |      ::|  |: : |  |      |||||| || :|
Db        597 LLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGFTPLHLAAQEGH 656

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
               |:  ||:::||    : |:|
Db        657 REMAALLIENGAKVGAQARNG---------------------------------------- 677

Qy        739 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 798
                        ||:|| |  : : ||| |::  | :: : | |  ||| |  :|
Db        678 --------------LTPMHLCAQEDRVSVAEELVKENAAIDPKTKAGYTPLHVACHFGQ 722

Qy        799 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 858
              :::   ||:: | |:   : ::||||:|||:|   :    || ||| | :     |||||
Db        723 INMVRFLIEHGARVSVITRASYTPLHQAAQQGHNSVVRYLLEHGASPNVHTSTGQTPLSI 782

Qy        859 VS----ADDVSALLT------AAMPPSALPSCYKPQVLNGVRSPGATADAL---SSGPSS 905
                   | || |          :     ||||      :| |  : :   |
Db        783 AERLGYVSVVEALKTITETTVITETTTVTEERYKPQ------NPEAMNETMFSDSEDEGE 836

Qy        906 PSSLSAASSLDNLSGSFSELSSVVSSSG------TEGASSLEKKEVPGVDFSITQFVRNL 959
               : ::| :   :  | | ::   :  |:|      ||   |    || | |  :   :|
Db        837 DNQITANAHAHDFSESLTK--GLHDSTGVHLIHATEPTLS-RSPEVEGTDGDLDALIRKA 893

Qy        960 GLEHL--------MDIFEREQITLDVLVEMGHKELKEIGINAYGH-----RHKLIKGVER 1006
               | :        :|    : :|: :   |     |    ::| |       ||    ||
Db        894 QHEPITTAMADPSLDASLPDNVTI-MRTTMQPSFLISFMVDARGGAMRGCRH---SGVRI 949

Qy       1007 LISGQQGLNP-YLTLNTSGSGTI---------------LIDLSPDDKEFQSV-------- 1042
              :|  ::   |  :|    |  :                    :::::|   :|
Db        950 IIPPRKAPQPTRVTCRYLGKDKLAHPPPLSEGEALASRILEMAPHGAKFLGPVILEVPHF 1009

Qy       1043 ----EEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 1098
                  | :  :   || |    :  : :   :   : | | :          :: |
Db       1010 ASLRGRREREIVILRSDDGQH----WKEHQLEATEDAVQEVLNESFDAEELSQLDDLHTSR 1065

Qy       1099 NERMLFHGSPFVNAIIHKGFDERH 1122
                |:| :  |   |:: :   | |
Db       1066 ITRILTNDFPMYFAVVTRVRQEVH 1089


RESULT 23
US-09-557-034-33
; Sequence 33, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Liang
;                   Blehm, E. Scot
;        TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                            PROTEINS, NUCLEIC ACID MOLECULES, AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 171
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Carol Talkington Verser, Ph.D.
;                        Heska Corporation
;             STREET: 1825 Sharp Point Drive
;             CITY: Fort Collins
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80525
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: WordPerfect for Windows, Version 7.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,034
;             FILING DATE: 21-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/065,474
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Verser, Carol Talkington
;             REGISTRATION NUMBER: 37,459
;             REFERENCE/DOCKET NUMBER: HW-5-C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 970/493-7272
;             TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 33:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1745 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-557-034-33

  Query Match           12.9%;  Score 835;  DB 4;  Length 1745;
  Best Local Similarity  25.7%;  Pred. No. 5e-63;
  Matches  314;  Conservative  154;  Mismatches  452;  Indels  304;  Gaps   37;

Qy         34 GSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSG---RRC-AGGGAACAS 89
              |  :|  | |:  | |              |:| ||          ||      | | |  |
Db         35 GESSASFLRAARAGNLD----------RVLELL--------RSGTDINTCNANGLNALHL 76

Qy         90 AAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRK-STPLHFAAGFGRKDVV 148
              |: |      ||| :  |  ||              | | || :| || |:  |:: :|
Db         77 ASKEGHHEVVRELLK--RKADV--------------DAATRKGNTALHIASLAGQELIV 119

Qy        149 EYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIK 208
                |::|||||  :   |  ||: |     |  ||  || | |:        :|||  |  :
Db        120 TVLVENGANVNVQSLNGFTPLYMAAQENHESVVRYLLAHNANQALSTEDGFTPLAVALQQ 179

Qy        209 GKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYKKDELLES----- 258
              |   |  |||::     :|       || :|       :|  :|   |:  |   :|
Db        180 GHDRVVAVLLENDTRGKVR----LPALHIAAKKDDTKAATLLLQNEHNSDVTSKSGFTPL 235

Qy        259 ---ARSGNEEKMMALL-------------TPLNV-------------------NCHAS 281
                 |  |||      ||              :||:|                    :|
Db        236 HIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTNMVSLLLAHGAVIDCRTR 295

Qy        282 DGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSY------------- 328
              |  |||| |:     ::| |||: || : || |  | |||  |
Db        296 D--LLTPLHCASRSGHDQVVDLLLEKGAPISAKTKNGLAPLHMAAQVDDVTVDYLTPLHV 353

Qy        329 ----GHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
                  ||  | :||:   |  ||  |  ||||| |  |||:::  ||| | |
Db        354 AAHCGHVRVAKLLLDRNADPNARALNGFTPLHIACKKNRIKIVELLLKYHA--------- 404

Qy        385 KSAIDLAPTPQLKERLAYEFKGH-----SLLQAAREADVTRIKKHLSLEMVNFKHPQTHE 439
                ||:       |  | |      |||   |||  ::                    |
Db        405 --AIEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVR---------------GE 447

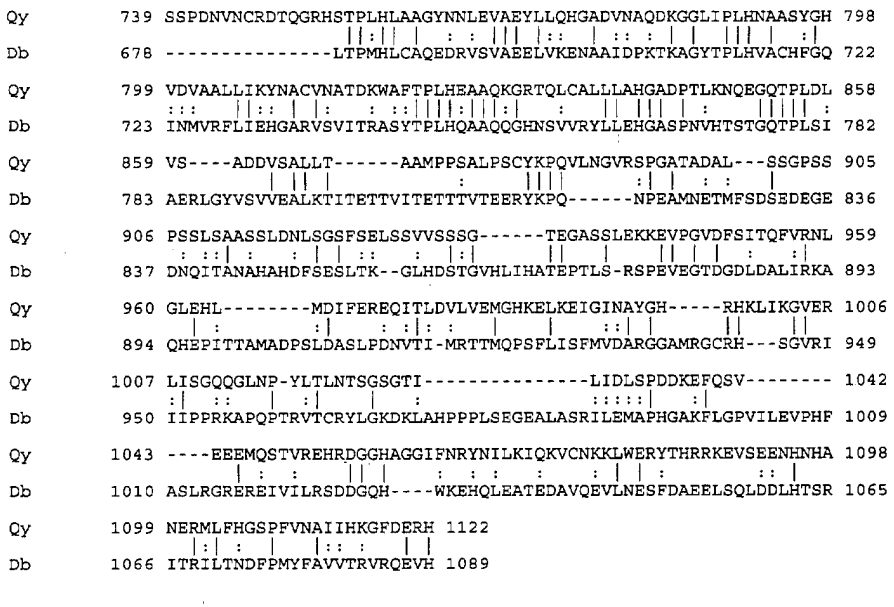
Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              | || || :    :  |  :|:| || ::   :|  ||||:||    : |:| ::::  |
Db        448 TPLHLAARA---NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANAS 504

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
               ||      | || ||  |  :   :|: :| |  ::: :||| | :
Db        505 PNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLHL---------------- 551

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVEY 618
                        ||| |::   | |  ::     ||||: | |||| || ||    |
Db        552 -------------AAKYGNLPVAKSL--LERGTPVDIEGKNQVTPLHVAAHYNNDKVALL 596

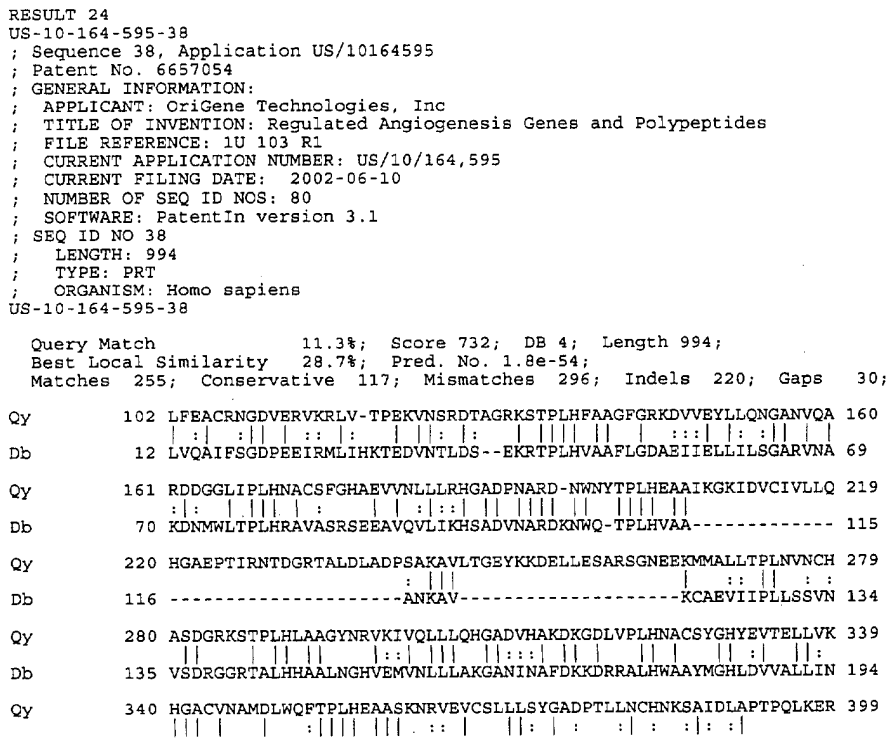
Qy        619 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 678
              ||::||  ||  | |  ||| |      ::|  |: : |  |      |||||| || :|
Db        597 LLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGFTPLHLAAQEGH 656

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
               |:  ||:::||    : |:|
Db        657 REMAALLIENGAKVGAQARNG---------------------------------------- 677
```



RESULT 24
US-10-164-595-38
; Sequence 38, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-38

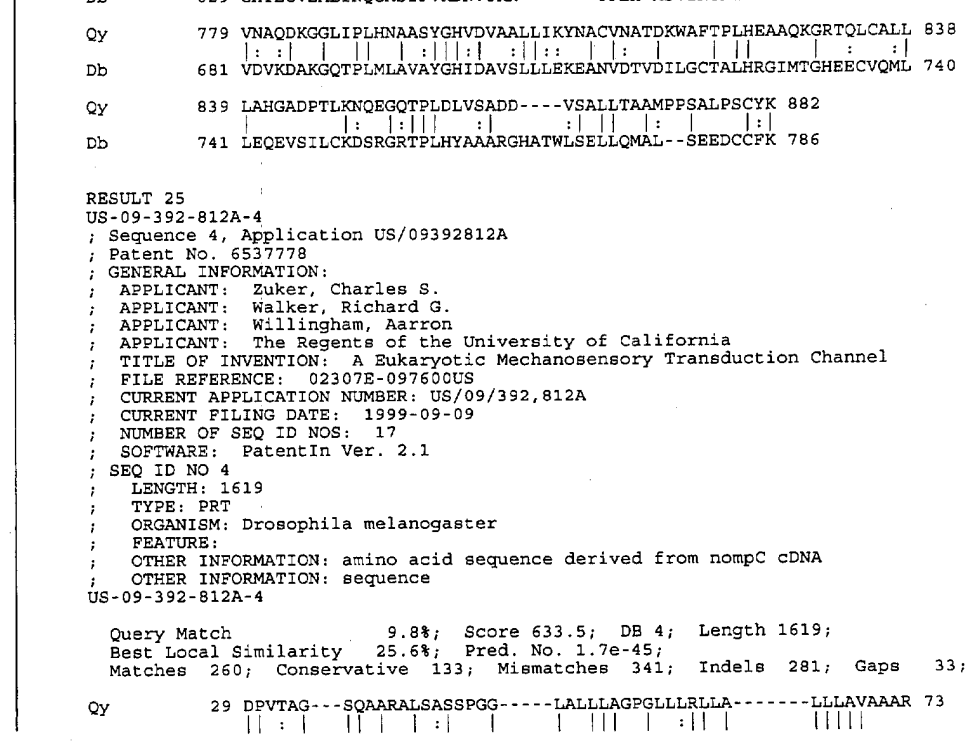
Query Match 11.3%; Score 732; DB 4; Length 994;
Best Local Similarity 28.7%; Pred. No. 1.8e-54;
Matches 255; Conservative 117; Mismatches 296; Indels 220; Gaps 30;

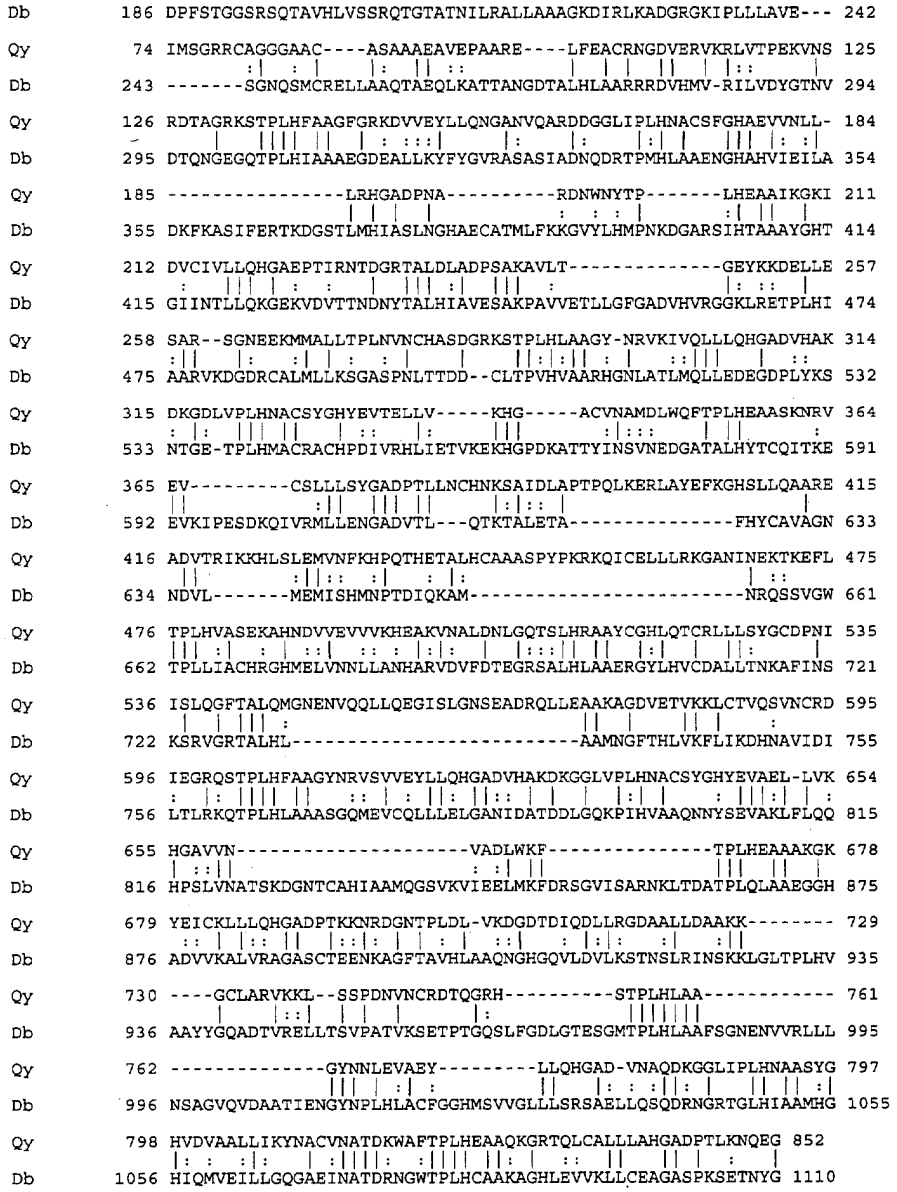




RESULT 25
US-09-392-812A-4
; Sequence 4, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/09/392,812A
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1619
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC cDNA
; OTHER INFORMATION: sequence
US-09-392-812A-4

Query Match 9.8%; Score 633.5; DB 4; Length 1619;
Best Local Similarity 25.6%; Pred. No. 1.7e-45;
Matches 260; Conservative 133; Mismatches 341; Indels 281; Gaps 33;

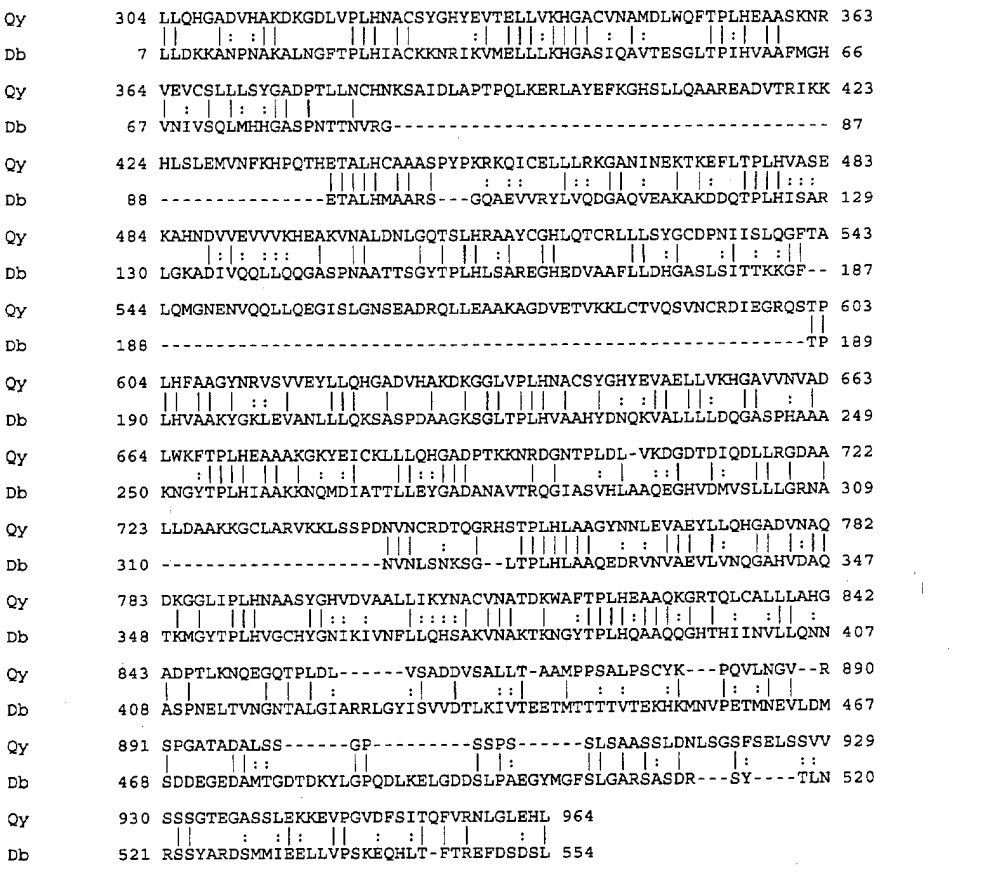




RESULT 26
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.

; APPLICANT: Devarajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identificat
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

Query Match 9.7%; Score 629.5; DB 3; Length 1088;
Best Local Similarity 28.1%; Pred. No. 1.9e-45;
Matches 195; Conservative 85; Mismatches 234; Indels 181; Gaps 16;



RESULT 27
US-10-164-595-71
; Sequence 71, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:

; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-71

Query Match 9.3%; Score 601.5; DB 4; Length 683;
Best Local Similarity 28.4%; Pred. No. 2.4e-43;
Matches 221; Conservative 103; Mismatches 314; Indels 139; Gaps 24;

```
Qy      134 TPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNA 193
Db        2 TAVH-----GRFTRSQTLIQNGGEIDCVDKDGNTPLHVAARYGHELLINTLITSGADTAK 56

Qy      194 RDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKD 253
Db       57 CGIHSMFPLHLAALNAHSDCCRKLLSSGQKYSI----------VSLFSNEHVLSAGFEID 106

Qy      254 ELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHA 313
Db      107 ----------------TP--------DKFGRTCLHAAAAGGNVECIKLLQSSGADFHK 140

Qy      314 KDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSY 373
Db      141 KDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDDWGRTALHYAAASD----------M 190

Qy      374 GADPTLL-NCHNKSAIDLAPTPQLKER---LAYEFKGHSLLQAAREADVTRIKKHLSLEM 429
Db      191 DRNKTILGNAHDNSE-ELERARELKEKEATLCLEF----LLQ----------------- 227

Qy      430 VNFKHPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT--PLHVASE 483
Db      228 -NDANPSIRDKEGYNSIHYAAAYGH---RQCLELLLERTNSGFEESDSGATKSPLHLAAY 283

Qy      484 KAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTA 543
Db      284 NGHHQALEVLLQSLVDLDIRDEKGRTALDLAAFKGHTECVEALIN----------QGASI 333

Qy      544 LQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTP 603
Db      334 FVKDNVTKRTPLHASVINGHTLCLRLLLEIAD----------NPEAVDVKDAKGQ--TP 380

Qy      604 LHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVAD 663
Db      381 LMLAVAYGHIDAVSLLLEKEANVDTVDILGCTALHRGIMTGHEECVQMLLEQEVSILCKD 440

Qy      664 LWKFTPLHEAAAKGKYEICKLLLQHG---ADPTKKNRDGNTPLD--LVKDGDTDIQDLLR 718
Db      441 SRGRTPLHYAAARGHATWLSELLQMALSEEDCCFKDNQGYTPLHWACYNGNENCIEVLLE 500

Qy      719 --------GD-------AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGY 763
Db      501 QKCFRKFIGNPFTPLHCAIINDHG--NCASLLLGAIDSSIVSCRDDKGR--TPLHAAAFA 556

Qy      764 NNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLI-KYNACVNATDKWAFTP 822
Db      557 DHVECLQLLLRHSAPVNAVDNSGKTALMMAAENGQAGAVDILVNSAQADLTVKDKDLNTP 616

Qy      823 LHEAAQKGRTQLCALLLAHGADPTL---KNQEGQTPLDLVSADDVSALLTAAMPPSA 876
Db      617 LHLACSKGHEKCALLILDKIQDESLINEKNNALQTPLHVAARNGLKVVVEELLAKGA 673
```

RESULT 28
US-09-392-812A-2
; Sequence 2, Application US/09392812A

; Patent No. 6537778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/09/392,812A
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-09-392-812A-2

Query Match 8.7%; Score 563; DB 4; Length 1704;
Best Local Similarity 24.8%; Pred. No. 2.7e-39;
Matches 210; Conservative 121; Mismatches 296; Indels 220; Gaps 22;

```
Qy      105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
Db        7 AAENGHAHVIEILADKFKASIFERTKDGSTLMHIASLNGHAECATMLFKKGVYLHMPNKD 66

Qy      165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
Db       67 GARSIHTAAAYGHTGIINTLLQKGEKVDVTTN-NYTALHIAVESAKPAVVETLLGFGADV 125

Qy      225 TIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR--SGNEEKMMALLTPLNVNCHASD 282
Db      126 HVRG------------------GKLRETPLHIAARVKDGDRCALMLLKSGASPNLTTDD 166

Qy      283 GRKSTPLHLAAGY-NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLV--- 338
Db      167 --CLTPVHVAARHGNLATLMQLLEDEGDPLYKSNTGE-TPLHMACRACHPDIVRHLIETV 223

Qy      339 --KHG-----ACVNAMDLWQFTPLHEAASKNRVEV---------CSLLLSYGADPTLLNC 382
Db      224 KEKHGPDKATTYINSVNEDGATALHYTCQITKEEVKIPESDKQIVRMLLENGADVTL--- 280

Qy      383 HNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETAL 442
Db      281 QTKTALETA---------------FHYCAVAGNNDVL-------MEMISHMNPTDIQKAM 318

Qy      443 HCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNA 502
Db      319 ------------------------NRQSSVGWTPLLIACHRGHMELVNNLLANHARVDV 353

Qy      503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLG 562
Db      354 FDTEGRSALHLAAERGYLHVCDALLTNKAFINSKSRVGRTALHL---------------- 397

Qy      563 NSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 622
Db      398 ----------AAMNGFTHLVKFLIKDHNAVIDILTLRKQTPLHLAAASGQMEVCQLLLEL 447

Qy      623 GADVHAKDKGGLVPLHNACSYGHYEVAEL-LVKHGAVVN---------------------- 660
Db      448 GANIDATDDLGQKPIHVAAQNNYSEVAKLFLQQHPSLVNATSKDGNTCAHIAAMQGSVKV 507

Qy      661 VADLWKF---------------TPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL 705
Db      508 IEELMKFDRSGVISARNKLTDATPLQLAAEGGHADVVKALVRAGASCTEENKAGFTAVHL 567

Qy      706 -VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYN 764
Db      568 AAQNGHGQVLDVLKSTNSLRINSKKLGL---------------------TPLHVAAYYG 605
```

```
Qy        765 NLE-VAEYLLQHGADVNAQ-----------DKGGLIPLHNAA---------------- 794
              :  | | |    | | ::           : |: ||| ||
Db        606 QADTVRELLTSVPATVKSETPTGQSLFGDLGTESGMTPLHLAAFSGNENVVRLLLNSAGV 665

Qy        795 ---------SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 845
                      :||: :  :|:   | :||||:  :|||| ||: |  ::  ||   ||| |
Db        666 QVDAATIENMHGHIQMVEILLGQGAEINATDRNGWTPLHCAAKAGHLEVVKLLCEAGASP 725

Qy        846 TLKNQEG 852
              :   |
Db        726 KSETNYG 732


RESULT 29
US-09-392-812A-6
; Sequence 6, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
;  APPLICANT:  Zuker, Charles S.
;  APPLICANT:  Walker, Richard G.
;  APPLICANT:  Willingham, Aarron
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION:  A Eukaryotic Mechanosensory Transduction Channel
;  FILE REFERENCE:  02307E-097600US
;  CURRENT APPLICATION NUMBER: US/09/392,812A
;  CURRENT FILING DATE:  1999-09-09
;  NUMBER OF SEQ ID NOS:  17
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 1709
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
;   FEATURE:
;   OTHER INFORMATION: amino acid sequence derived from nompC genomic
;   OTHER INFORMATION: sequence
US-09-392-812A-6

  Query Match             8.4%;  Score 545.5;  DB 4;  Length 1709;
  Best Local Similarity   24.5%;  Pred. No. 9.1e-38;
  Matches  260;  Conservative  114;  Mismatches  352;  Indels  335;  Gaps   38;

Qy         64 ALLLAVAAAR-------IMSGRRC-----AGGGA---ACASAAAEAVEPAARELFEACRN 108
              | |||| | :       |  | |       |  |    |   :  | :|   :   |  |
Db        154 AFLLAVKAGKDQIVDKMIRKGARVDYSTKDGRNATHIAAMYSGVETLELILKRYSELLRK 213

Qy        109 GDVERVKRLVTPEKVNSRDTAG-RKSTPLHFAAGFGRK---DVVEYLLQNGANVQARDDG 164
              |              || :|   :| |     |    :|: :|::       | |
Db        214 G------------------AGPKKQLAIHVACERKSKKAFPIVKRILEDTDQRMAEDGD 254

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNAR--DNWNYTPLHEAAIKGKIDVCIVLLQHGA 222
              | :|:| |  ||:  :| |||    :|   |  |     | || ||  | |:     :  |
Db        255 GSLPIHLAFKFGNVNIVELLLSGPSDEQTRKADGNGDTLLHLAARSGNIEAVRTAIAAGC 314

Qy        223 E-PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHAS 281
              :   ::|  ||| |           :|| |              :: |: ::  |  : :
Db        315 DNANVQNRVGRTPLHEC-----LTVTGTQK-----GYVAEVGDQNMLKIMFKLRADANIH 364

Qy        282 DGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAK------------------------- 314
              |    ||:|:||      :|: |: : |  : |:
Db        365 DKEDKTPVHVAAERGDTSMVESLIDKFGGSIRARTRDGSTLLHIAACSGHTSTALAFLKR 424

Qy        315 ------DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCS 368
                   :|   : ||:| : |  :| ::|:  |  |:       :| || |    :  |
Db        425 VPLFMPNKKGALGLHSAAAAGFNDVVKMLIARGTNVDVRTRDNYTALHVAVQSGKASVVE 484

Qy        369 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE 428
              ||  |||                                ||  |:
Db        485 TLLGSGAD-------------------------IHVKGGELMDG---------------- 503

Qy        429 MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHND 488
                  || || ||  |     | |  ||| : |:    :|   ||| ||::  : :
Db        504 ----------ETCLHIAARS---GNKDIM-LLLDENADSKISSKIGETPLQVAAKSCNFE 549

Qy        489 VVEVVVKH----------EAKVNALDNLGQTSLHRAAYCGHLQ---------TCRLLLSY 529
                 :::||          :  ||    | | |:|| ||      |              ||: |
Db        550 AASMILKHLSEVLTQEQLKEHVNHRTNDGFTALHYAAEIEQRQLHFPGEDAKLVNLLIDY 609

Qy        530 GCDPNIISLQ-GFTALQM----GNENVQQLLQEGISLG------NSEADR---QLLEAAK 575
              |    : ||    ||: |    ||: |   :   | |       | ::      ||||
Db        610 GGMVEMPSLNANETAMHMAARSGNQAVLLAMVNKIGAGAVQIVQNKQSKNGWSPLLEACA 669

Qy        576 AGD------VETVKKLC---------------------------------------- 586
              |       :  | ||
Db        670 RGHSGVANILLKVLVLCVGPGPGPGPRLQGRGYWTRTRARVTVPWLQYQGYWARTRTRTR 729

Qy        587 ----------------------TV--------QSVNCRDIEGRQSTPLHFAAGYNRVSV 615
                               ||          ::  |  ||  | || ||    :|:
Db        730 ARATGPGLQDQGYWARTRTRTKVTVPRLLGDHHARIDVFDEMGR--TALHLAAFNGHLSL 787

Qy        616 VEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVK-HGAVVNVADLWKFTPLHEAA 674
              |  |||| | |::| |  |  ||| |  :|| :|  :||: ||| :    |  | || ||
Db        788 VHLLLQHKAFVNSKSKTGEAPLHLAAQHGHVKVVNVLVQDHGAALEAITLDNQTALHFAA 847

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD-TDIQDLL-------RGDAALLD- 725
                |:  : : ||  ||:|  ::  | ||| |  : |   |:  |         |     :|
Db        848 KFGQLAVSQTLLALGANPNARDDKGQTPLHLAAENDFPDVVKLFLKMRNNNRSVLTAIDH 907

Qy        726 --------AAKKGCLARVKKLSSPDN--VNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQH 775
                     || || || |::|   |   |     |:   :| ||:||   :  : : ||::
Db        908 NGFTCAHIAAMKGSLAVVRELMMIDKPMVIQAKTKTLEATTLHMAAAGGHANIVKILLEN 967

Qy        776 GADVNAQD---------------------------------KGGLIPLHNAASYGHVDVA 802
              ||  ||:|                                 | ||  || || ||: |
Db        968 GA--NAEDENSGMTALHLGAKNGFISILEAFDKILWKRCSRKTGLNALHIAAFYGNSDFV 1025

Qy        803 ALLIK------------YNACVNA--TDKWAFTPLHEAAQKGRTQLCALLLAHG--ADPT 846
                ::|             ||  ||   : :: |||| ||  |   |  :||  |   | |
Db       1026 NEMLKHVQATVRSEPPIYNHHVNKEFSTEYGFTPLHLAAHSGHDSLVRMLLNQGVQVDAT 1085

Qy        847 ----LKNQEGQTPLDLVSAD------DVSALLTAAMPPSAL 877
                  :  :| :   ||::       |  |:||    :||
Db       1086 STTMMSEKEKERAKDLLNVAVFSENMAVELLITATEYNAAL 1126


RESULT 30
US-09-065-474-139
; Sequence 139, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  139:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-09-065-474-139

  Query Match             8.0%;  Score 514.5;  DB 3;  Length 352;
  Best Local Similarity   23.7%;  Pred. No. 3e-36;
  Matches  137;  Conservative   64;  Mismatches  139;  Indels  237;  Gaps    4;

Qy        131 RKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGAD 190
              |  |||| ||   : |:|  |::||| | |      ||| |   |: ::| |||:  |
Db         10 RGETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANAS 69

Qy        191 PNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEY 250
              |||     ||||| || :|: :|  :|: || : |:    |
Db         70 PNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKG------------------ 110

Qy        251 KKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGAD 310
                                                 ||||||| |  : : : ||: |
Db        111 ----------------------------------FTPLHLAAKYGNLPVAKSLLERGTP 135

Qy        311 VHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLL 370
              |  : |  : ||| |  | : :|  ||:::||  :|     :|||| || ||:::: | |
Db        136 VDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTL 195

Qy        371 LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMV 430
              | |
Db        196 LHY------------------------------------------------------- 198

Qy        431 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 490
                                              || | ::|   ||||:|::: | ::
Db        199 -------------------------------KANANAESKAGFTPLHLAAQEGHREMA 225

Qy        491 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 550
               ::::: ||| |
Db        226 ALLIENGAKVGA---------------------------------------------- 237

Qy        551 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 610
                                    |: |                          ||:|  |
Db        238 ---------------------QARNG----------------------LTPMHLCAQE 252

Qy        611 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 670
              :|||| | |::  | :  | | |  ||| || :|   :   |::||| |:|     :|||
Db        253 DRVSVAEELVKENAAIDPKTKAGYTPLHVACHFGQINMVRFLIEHGARVSVITRASYTPL 312

Qy        671 HEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
              |:|| :|   : : ||:||| |      | ||| : :
Db        313 HQAAQQGHNSVVRYLLEHGASPNVHTSTGQTPLSIAE 349


RESULT 31
US-09-557-034-139
; Sequence 139, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 139:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 352 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: Protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 139:
US-09-557-034-139

  Query Match             8.0%;  Score 514.5;  DB 4;  Length 352;
  Best Local Similarity   23.7%;  Pred. No. 3e-36;
  Matches  137;  Conservative   64;  Mismatches  139;  Indels  237;  Gaps    4;

Qy        131 RKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGAD 190
              |  |||| ||   : |:|  |::||| | |      ||| |   |: ::| |||:  |
Db         10 RGETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANAS 69

Qy        191 PNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEY 250
              |||     ||||| || :|: :|  :|: || : |:    |
Db         70 PNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKG------------------ 110

Qy        251 KKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGAD 310
                                                 ||||||| |  : : : ||: |
Db        111 ----------------------------------FTPLHLAAKYGNLPVAKSLLERGTP 135

Qy        311 VHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLL 370
              |  : |  : ||| |  | : :|  ||:::||  :|     :|||| || ||:::: | |
Db        136 VDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTL 195

Qy        371 LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMV 430
              | |
Db        196 LHY------------------------------------------------------- 198

Qy        431 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 490
                                              || | ::|   ||||:|::: | ::
Db        199 -------------------------------KANANAESKAGFTPLHLAAQEGHREMA 225

Qy        491 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 550
               ::::: ||| |
Db        226 ALLIENGAKVGA---------------------------------------------- 237

Qy        551 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 610
                                    |: |                          ||:|  |
Db        238 ---------------------QARNG----------------------LTPMHLCAQE 252

Qy        611 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 670
              :|||| | |::  | :  | | |  ||| || :|   :   |::||| |:|     :|||
Db        253 DRVSVAEELVKENAAIDPKTKAGYTPLHVACHFGQINMVRFLIEHGARVSVITRASYTPL 312

Qy        671 HEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
              |:|| :|   : : ||:||| |      | ||| : :
Db        313 HQAAQQGHNSVVRYLLEHGASPNVHTSTGQTPLSIAE 349
```

```
RESULT 32
US-09-972-115A-22
; Sequence 22, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-22

  Query Match             6.8%;  Score 440;  DB 4;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 9e-31;
  Matches   90;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        869 TAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSV 928
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSV 60

Qy        929 VSSSGTEGASSLEKKEVPGVDFSITQFVRN 958
              ||||||||||||||||||||||||||||||
Db         61 VSSSGTEGASSLEKKEVPGVDFSITQFVRN 90


RESULT 33
US-09-031-485-23
; Sequence 23, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-031-485-23

  Query Match             6.5%;  Score 421.5;  DB 2;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 3e-28;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 526
              ::: | ::||||||                                  ||:|||::  :||
Db          1 VDDVTVDYLTPLHV--------------------------------AAHCGHVRVAKLL 27

Qy        527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
              |    |||  :| ||| | :                            | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI------------------------ACKKNRIKIVELLL 61

Qy        587 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 643
                  :      ||       :||| ||    :::| |||| ||:        |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        644 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
                ::  :||::||  |: |     |||| |:  |  :|  |||  | |    ||  |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        704 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 759
              :  |:|  ::  :|     |  ||   |||                          |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        760 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 819
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        820 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 850
              :|||| ||:| :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 34
US-08-847-429A-23
; Sequence 23, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
```

```
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-429A-23

  Query Match             6.5%;  Score 421.5;  DB 2;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 3e-28;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 526
              ::: | ::||||||                              ||:|||::  :||
Db          1 VDDVTVDYLTPLHV--------------------------------AAHCGHVRVAKLL 27

Qy        527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
              |   |||  :| ||| | :                            | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI--------------------------ACKKNRIKIVELLL 61

Qy        587 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 643
              :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        644 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
              ::  :||::|| |: |     |||| |:  |  :|  |||| | |    ||  |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        704 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 759
              :  |:|  ::  :|     |  ||   |||                        |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        760 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 819
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        820 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 850
              :|||| ||:| :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 35
US-09-065-474-23
; Sequence 23, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-065-474-23

  Query Match             6.5%;  Score 421.5;  DB 3;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 3e-28;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 526
              ::: | ::||||||                              ||:|||::  :||
Db          1 VDDVTVDYLTPLHV--------------------------------AAHCGHVRVAKLL 27

Qy        527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
              |   |||  :| ||| | :                            | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI--------------------------ACKKNRIKIVELLL 61

Qy        587 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 643
              :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        644 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
              ::  :||::|| |: |     |||| |:  |  :|  |||| | |    ||  |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        704 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 759
              :  |:|  ::  :|     |  ||   |||                        |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        760 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 819
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        820 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 850
              :|||| ||:| :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 36
US-09-557-034-23
; Sequence 23, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
```

```
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,034
;             FILING DATE: 21-Apr-2000
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/065,474
;             FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Verser, Carol Talkington
;             REGISTRATION NUMBER: 37,459
;             REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 970/493-7272
;             TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 23:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 303 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-557-034-23

  Query Match             6.5%;  Score 421.5;  DB 4;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 3e-28;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 526
              ::: | ::||||||                                  ||:|||::  :||
Db          1 VDDVTVDYLTPLHV--------------------------------AAHCGHVRVAKLL 27

Qy        527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
              |   |||  :| ||| | :                  | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI------------------------ACKKNRIKIVELLL 61

Qy        587 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 643
                :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        644 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
                ::  :||::|| |: |     |||| |:  |  :|  |||| | |     ||  |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        704 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 759
              :  |:|  ::  :|      |  ||   |||                      |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        760 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 819
              || | || ||: ||:  |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        820 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 850
              :|||| ||:|  :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 37
US-09-031-485-38
; Sequence 38, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  302 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-031-485-38

  Query Match             6.3%;  Score 409;  DB 2;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 3.7e-27;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |     | : || ||: | :      ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------ 106

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619
                                                  |||| ||  |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        620 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 679
              :::|| | |  :   ||| |   |: ::  ||::  |  | |     :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        680 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              |:  :|: ||:|  |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        739 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 797
                        | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        798 HVDVAALLIKYNACVNATDK 817
              :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302


RESULT 38
US-08-847-429A-38
; Sequence 38, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
```

```
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  302 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-429A-38

  Query Match             6.3%;  Score 409;  DB 2;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 3.7e-27;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |    | : || ||: | :     ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------- 106

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619
                                                  |||| ||   |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        620 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 679
              :::|| | |  :    ||| |   |: ::  ||::  |  | |   :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        680 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              |:  :|: ||:|  |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        739 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 797
                      | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        798 HVDVAALLIKYNACVNATDK 817
               :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302


RESULT 39
US-09-065-474-38
; Sequence 38, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  302 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-065-474-38

  Query Match             6.3%;  Score 409;  DB 3;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 3.7e-27;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |    | : || ||: | :     ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------- 106

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619
                                                  |||| ||   |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        620 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 679
              :::|| | |  :    ||| |   |: ::  ||::  |  | |   :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        680 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              |:  :|: ||:|  |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        739 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 797
                      | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        798 HVDVAALLIKYNACVNATDK 817
               :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302


RESULT 40
US-09-557-034-38
```

; Sequence 38, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 302 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-557-034-38

Query Match 6.3%; Score 409; DB 4; Length 302;
Best Local Similarity 29.5%; Pred. No. 3.7e-27;
Matches 112; Conservative 56; Mismatches 124; Indels 88; Gaps 6;

```
Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              |  ||  ||   :  :   : :||| : |: | :    ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 559
              : |    | : || ||: | :     ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------- 106

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619
                                                 |||| ||  |:  :| |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        620 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 679
              :::|| | |  :    ||| |   |: ::  ||::  |  | |    :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        680 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 738
              |:  :|: ||:| |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        739 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 797
                        | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282
```

```
Qy        798 HVDVAALLIKYNACVNATDK 817
              :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302
```

RESULT 41
US-09-439-313-378
; Sequence 378, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-439-313-378

Query Match 6.2%; Score 398.5; DB 4; Length 1719;
Best Local Similarity 21.3%; Pred. No. 6.3e-25;
Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;

```
Qy         80 CAGGGAACASAAAEAVEPAAREL-----------FEACRNGDVERV--------KRLVTP 120
              | | | :   |: :  : | : |           |  ||       :|         : |
Db         72 CRGSGKSNVGASGDHDDSAMKTLRNKMGKWCCHCFPCCRGSGKSKVGAWGDYDDSAFMEP 131

Qy        121 EKVNSRDTAGRKSTPLHFAAGFG---RKDVVEYLLQNGANVQARDDGGLIPLHNACSFGH 177
                       |      || || :|   |||::  ::     :|  :|       || | : |:
Db        132 R----YHVRGEDLDKLHRAAWWGKVPRKDLI--VMLRDTDVNKKDKQKRTALHLASANGN 185

Qy        178 AEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDL 237
              :||| |||       |  ||    | | :|    : :  ::||:|| :| | :  | | |
Db        186 SEVVKLLLDRRCQLNVLDNKKRTALIKAVQCQEDECALMLLEHGTDPNIPDEYGNTTLHY 245

Qy        238 ADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNR 297
              |                                ||:|:||
Db        246 A----------------------IYNEDKLMA---------------------------- 255

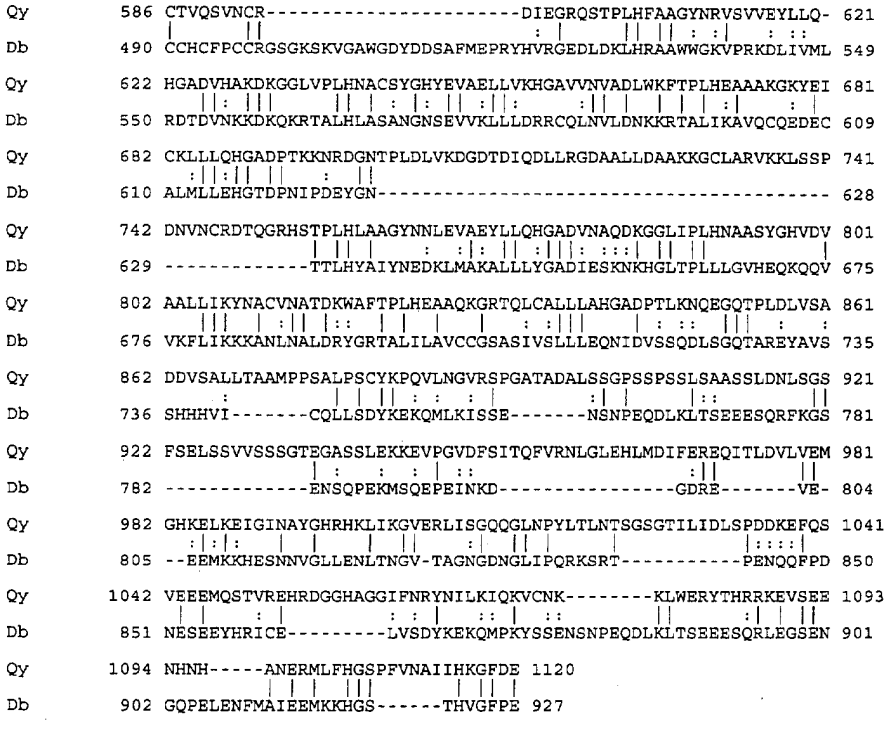
Qy        298 VKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHE 357
                 : || :|||: :|:|  | ||       :| : |:|  | :|||:| :  | |
Db        256 ----KALLLYGADIESKNKHGLTPLLLGVHEQKQQVVKFLIKKKANLNALDRYGRTALIL 311

Qy        358 AASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTP--QLKERLAYEFKGHSLLQAARE 415
              |       : ||||    | :  :   ::| : | :    :  :|  ::|   :|: : |
Db        312 AVCCGSASIVSLLLEQNIDVSSQDLSGQTAREYAVSSHHHVICQLLSDYKEKQMLKISSE 371

Qy        416 ----ADVTRIKKHLSLEMVNFKHPQTH---ETALHCAAAS---PYPKRKQICELLLRKGA 465
                  :|:| :            |:||    |      ||:|   |:  | :: :    |
Db        372 NSNPENVSRTR----------NKPRTHMVVEVDSMPAASSVKKPFGLRSKMGKWCCRCFP 421

Qy        466 NINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL 525
              | |      :|:  |:|          |          :|:       :| |   |
Db        422 CCRESGKS-----NVGTSGDHDD--------SAMKTLRSKMGK-------WCRHCFPC-- 459

Qy        526 LLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKL 585
              |  :           |  ||       | |  : ::  :  |           : |
Db        460 -----CRGS------------GKSNV------GASGDHDDSAMKTLR-------NKMGKW 489
```



RESULT 42
US-09-352-616A-378
; Sequence 378, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqui
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 378
; LENGTH: 1719
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-352-616A-378

Query Match 6.2%; Score 398.5; DB 4; Length 1719;
Best Local Similarity 21.3%; Pred. No. 6.3e-25;
Matches 237; Conservative 137; Mismatches 411; Indels 327; Gaps 34;





```
RESULT 43
US-09-636-215-378
; Sequence 378, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE:  210121.42717C17
;  CURRENT APPLICATION NUMBER: US/09/636,215
;  CURRENT FILING DATE:  2000-08-10
;  NUMBER OF SEQ ID NOS:  852
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 378
;   LENGTH: 1719
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-636-215-378

  Query Match             6.2%;  Score 398.5;  DB 4;  Length 1719;
  Best Local Similarity   21.3%;  Pred. No. 6.3e-25;
  Matches  237;  Conservative  137;  Mismatches  411;  Indels  327;  Gaps   34;

Qy          80 CAGGGAACASAAAEAVEPAAREL-----------FEACRNGDVERV--------KRLVTP 120
Db          72 CRGSGKSNVGASGDHDDSAMKTLRNKMGKWCCHCFPCCRGSGKSKVGAWGDYDDSAFMEP 131

Qy         121 EKVNSRDTAGRKSTPLHFAAGFG---RKDVVEYLLQNGANVQARDDGGLIPLHNACSPGH 177
Db         132 R----YHVRGEDLDKLHRAAWWGKVPRKDLI--VMLRDTDVNKKDKQKRTALHLASANGN 185

Qy         178 AEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDL 237
Db         186 SEVVKLLLDRRCQLNVLDNKKRTALIKAVQCQEDECALMLLEHGTDPNIPDEYGNTTLHY 245

Qy         238 ADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNR 297
Db         246 A----------------------IYNEDKLMA---------------------------- 255

Qy         298 VKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHE 357
Db         256 ----KALLLYGADIESKNKHGLTPLLLGVHEQKQQVVKFLIKKKANLNALDRYGRTALIL 311

Qy         358 AASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTP--QLKERLAYEFKGHSLLQAARE 415
Db         312 AVCCGSASIVSLLLEQNIDVSSQDLSGQTAREYAVSSHHHVICQLLSDYKEKQMLKISSE 371

Qy         416 ----ADVTRIKKHLSLEMVNFKHPQTH---ETALHCAAAS---PYPKRKQICELLLRKGA 465
Db         372 NSNPENVSRTR----------NKPRTHMVVEVDSMPAASSVKKPFGLRSKMGKWCCRCFP 421

Qy         466 NINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL 525
Db         422 CCRESGKS-----NVGTSGDHDD--------SAMKTLRSKMGK-------WCRHCFPC-- 459

Qy         526 LLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKL 585

Db         460 -----CRGS------------GKSNV------GASGDHDDSAMKTLR-------NKMGKW 489

Qy         586 CTVQSVNCR----------------------DIEGRQSTPLHFAAGYNRVSVVEYLLQ- 621
Db         490 CCHCFPCCRGSGKSKVGAWGDYDDSAFMEPRYHVRGEDLDKLHRAAWWGKVPRKDLIVML 549

Qy         622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
Db         550 RDTDVNKKDKQKRTALHLASANGNSEVVKLLLDRRCQLNVLDNKKRTALIKAVQCQEDEC 609

Qy         682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
Db         610 ALMLLEHGTDPNIPDEYGN---------------------------------------- 628

Qy         742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
Db         629 ------------TTLHYAIYNEDKLMAKALLLYGADIESKNKHGLTPLLLGVHEQKQQV 675

Qy         802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
Db         676 VKFLIKKKANLNALDRYGRTALILAVCCGSASIVSLLLEQNIDVSSQDLSGQTAREYAVS 735

Qy         862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
Db         736 SHHHVI-------CQLLSDYKEKQMLKISSE-------NSNPEQDLKLTSEEESQRFKGS 781

Qy         922 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 981
Db         782 ------------ENSQPEKMSQEPEINKD----------------GDRE-------VE- 804

Qy         982 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1041
Db         805 --EEMKKHESNNVGLLENLTNGV-TAGNGDNGLIPQRKSRT-----------PENQQFPD 850

Qy        1042 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK--------KLWERYTHRRKEVSEE 1093
Db         851 NESEEYHRICE---------LVSDYKEKQMPKYSSENSNPEQDLKLTSEEESQRLEGSEN 901

Qy        1094 NHNH-----ANERMLFHGSPFVNAIIHKGFDE 1120
Db         902 GQPELENFMAIEEMKKHGS------THVGFPE 927

RESULT 44
US-09-685-166A-378
; Sequence 378, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  Harlocker, Susan L.
;  APPLICANT:  Jiang, Yuqui
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Day, Craig H.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Li, Samuel
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Hepler, William
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF PROSTATE CANCER
;  FILE REFERENCE: 210121.427C21
;  CURRENT APPLICATION NUMBER: US/09/685,166A
;  CURRENT FILING DATE:  2000-10-10
;  NUMBER OF SEQ ID NOS: 898
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 378
```

```
;   LENGTH: 1719
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-685-166A-378

  Query Match             6.2%;  Score 398.5;  DB 4;  Length 1719;
  Best Local Similarity   21.3%;  Pred. No. 6.3e-25;
  Matches  237;  Conservative  137;  Mismatches  411;  Indels  327;  Gaps   34;

Qy          80 CAGGGAACASAAAEAVEPAAREL-----------FEACRNGDVERV--------KRLVTP 120
               | | | :   |: :  : | :            |  ||      :|           : |
Db          72 CRGSGKSNVGASGDHDDSAMKTLRNKMGKWCCHCFPCCRGSGKSKVGAWGDYDDSAFMEP 131

Qy         121 EKVNSRDTAGRKSTPLHFAAGFG---RKDVVEYLLQNGANVQARDDGGLIPLHNACSFGH 177
                        |      || || :|   |||::  ::    :|  :|      || | : |:
Db         132 R----YHVRGEDLDKLHRAAWWGKVPRKDLI--VMLRDTDVNKKDKQKRTALHLASANGN 185

Qy         178 AEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDL 237
               :||| |||       |  ||   | | :|    : :  ::||:|| :| | :  | | |
Db         186 SEVVKLLLDRRCQLNVLDNKKRTALIKAVQCQEDECALMLLEHGTDPNIPDEYGNTTLHY 245

Qy         238 ADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNR 297
                  |                        ||:|:||
Db         246 A----------------------IYNEDKLMA---------------------------- 255

Qy         298 VKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHE 357
                 : || :|||: :|:|  | ||          :| : |:|  | :||:| :  | |
Db         256 ----KALLLYGADIESKNKHGLTPLLLGVHEQKQQVVKFLIKKKANLNALDRYGRTALIL 311

Qy         358 AASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTP--QLKERLAYEFKGHSLLQAARE 415
               |      : ||||    | :  :   ::| : | :    :  :|  ::|   :|: : |
Db         312 AVCCGSASIVSLLLEQNIDVSSQDLSGQTAREYAVSSHHHVICQLLSDYKEKQMLKISSE 371

Qy         416 ----ADVTRIKKHLSLEMVNFKHPQTH---ETALHCAAAS---PYPKRKQICELLLRKGA 465
                   :|:| :            |:||   |      ||:|   |:  | :: :   |
Db         372 NSNPENVSRTR----------NKPRTHMVVEVDSMPAASSVKKPFGLRSKMGKWCCRCFP 421

Qy         466 NINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL 525
                 |  |      :| :   |:|          |       :|:       :| |   |
Db         422 CCRESGKS-----NVGTSGDHDD--------SAMKTLRSKMGK-------WCRHCFPC-- 459

Qy         526 LLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKL 585
                 | :           | ||     | |  : ::  : |           : |
Db         460 -----CRGS-----------GKSNV------GASGDHDDSAMKTLR-------NKMGKW 489

Qy         586 CTVQSVNCR----------------------DIEGRQSTPLHFAAGYNRVSVVEYLLQ- 621
               |      ||                      : |     || || : :|   : ::
Db         490 CCHCFPCCRGSGKSKVGAWGDYDDSAFMEPRYHVRGEDLDKLHRAAWWGKVPRKDLIVML 549

Qy         622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
               ||: |||     || | : |: || :||:     :|| |  | | | :|    : |
Db         550 RDTDVNKKDKQKRTALHLASANGNSEVVKLLLDRRCQLNVLDNKKRTALIKAVQCQEDEC 609

Qy         682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
                 :||:|| ||    :  ||
Db         610 ALMLLEHGTDPNIPDEYGN---------------------------------------- 628

Qy         742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
                         | ||  |     :  :|: || :|||: :::| || ||          |
Db         629 ------------TTLHYAIYNEDKLMAKALLLYGADIESKNKHGLTPLLLGVHEQKQQV 675

Qy         802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
                  |||  | :||  ||:  |  |    |   : :|||   | : ::  |||  :   :
Db         676 VKFLIKKKANLNALDRYGRTALILAVCCGSASIVSLLLEQNIDVSSQDLSGQTAREYAVS 735

Qy         862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
                     :         | | ||  | :  :  : |        :| |     |::       ||
Db         736 SHHHVI-------CQLLSDYKEKQMLKISSE-------NSNPEQDLKLTSEEESQRFKGS 781

Qy         922 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 981
                        | :   :  : | ::                    :||       ||
Db         782 ------------ENSQPEKMSQEPEINKD---------------GDRE-------VE- 804

Qy         982 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1041
                 :|:|:   |  |    |  ||    :|  || |      |           |::::|
Db         805 --EEMKKHESNNVGLLENLTNGV-TAGNGDNGLIPQRKSRT----------PENQQFPD 850

Qy        1042 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK--------KLWERYTHRRKEVSEE 1093
                 | |    : |         : :  |   ::  |  ::          ||       :| | ||
Db         851 NESEEYHRICE---------LVSDYKEKQMPKYSSENSNPEQDLKLTSEEESQRLEGSEN 901

Qy        1094 NHNH-----ANERMLFHGSPFVNAIIHKGFDE 1120
                      | | |  |||       | || |
Db         902 GQPELENFMAIEEMKKHGS------THVGFPE 927


RESULT 45
US-09-031-485-28
; Sequence 28, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-031-485-28

  Query Match             6.1%;  Score 392;  DB 2;  Length 348;
  Best Local Similarity   26.5%;  Pred. No. 1.5e-25;
  Matches  116;  Conservative   56;  Mismatches  142;  Indels  124;  Gaps    6;

Qy         431 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 490
                 | :|    |::     |:      :: | ||| | :||       |  ||:||:: |::||
Db          28 NSQHSNKGESSASFLRAARAGNLDRVLE-LLRSGTDINTCNANGLNALHLASKEGHHEVV 86

Qy         491 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 550
                   ::| :| :|    | |:|| |:  |      :|:  | : |: || ||
Db          87 RELLKRKADVDAATRKGNTALHIASLAGQELIVTVLVENGANVNVQSLNGF--------- 137

Qy         551 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 610
                                                          |||: ||
```

```
Db        138 ----------------------------------------------------TPLYMAAQE 146

Qy        611 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 670
              |  ||| ||| | |:     : |  ||  |   ||  |  :|:::     |    :   |
Db        147 NHESVVRYLLAHNANQALSTEDGFTPLAVALQQGHDRVVAVLLENDTRGKV----RLPAL 202

Qy        671 HEAAAKGKYEICKLLLQ--HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 728
              | || |   :   |||  | :| | |:
Db        203 HIAAKKDDTKAATLLLQNEHNSDVTSKS------------------------------- 230

Qy        729 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 788
                                          ||||:|| | |  ||: ||: ||:|| | :  :
Db        231 ----------------------GFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNIS 266

Qy        789 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 848
              ||| |  :|  :: :||: : | ::   :   |||| |:: |  |:  |||  ||  : |
Db        267 PLHVATKWGRTNMVSLLLAHGAVIDCRTRDLLTPLHCASRSGHDQVVDLLLEKGAPISAK 326

Qy        849 NQEGQTPLDLVS-ADDVS 865
               : |  || : :  |||:
Db        327 TKNGLAPLHMAAQVDDVT 344
```

Search completed: March 25, 2004, 16:45:53
Job time : 35.2564 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:36:25 ; Search time 58.8205 Seconds
(without alignments)
5956.407 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6464	100.0	1240	4	AAY97749	Aay97749 Tankyrase
2	6464	100.0	1240	6	ABG73731	Abg73731 Tankyrase
3	6375.5	98.6	1262	4	AAB66290	Aab66290 Human tan
4	6375.5	98.6	1385	4	AAB66294	Aab66294 Human tan
5	6121	94.7	1169	4	AAB66288	Aab66288 Human tan
6	6121	94.7	1169	4	AAB66278	Aab66278 Human tan
7	6108	94.5	1166	4	AAB66295	Aab66295 Human tan
8	6092	94.2	1166	3	AAB27211	Aab27211 Human tan
9	6088	94.2	1166	5	ABP69457	Abp69457 Human pol
10	6053	93.6	1166	4	AAY72589	Aay72589 Human tan
11	5766	89.2	1100	4	AAY97748	Aay97748 Tankyrase
12	5766	89.2	1100	6	ABG73732	Abg73732 Tankyrase
13	5766	89.2	1100	6	ABG73730	Abg73730 Tankyrase
14	5760	89.1	1100	6	ABG73735	Abg73735 Tankyrase
15	5658	87.5	1084	6	ABG73736	Abg73736 Tankyrase
16	5605	86.7	1074	2	AAY05734	Aay05734 Human Grb
17	5543	85.8	1166	4	AAB47022	Aab47022 Human SPA
18	5474	84.7	1431	3	AAB27210	Aab27210 Human tan
19	5303.5	82.0	1083	6	AAE35352	Aae35352 Human col
20	5103.5	79.0	1327	3	AAY44402	Aay44402 Human tan
21	5103.5	79.0	1327	3	AAB27212	Aab27212 Human tan
22	5103.5	79.0	1327	4	AAB66279	Aab66279 Human tan
23	5103.5	79.0	1327	5	AAU79537	Aau79537 Human tan
24	5040.5	78.0	1333	3	AAB27209	Aab27209 Human tan
25	5038.5	77.9	991	4	AAB47023	Aab47023 Mouse SPA
26	4624.5	71.5	907	4	AAB48574	Aab48574 Human bre
27	4102	63.5	784	4	AAB66285	Aab66285 Human tan
28	4005	62.0	1181	4	ABB60894	Abb60894 Drosophil
29	3997	61.8	1181	4	AAB66297	Aab66297 Drosophil
30	3959	61.2	756	4	AAB66286	Aab66286 Human tan
31	3938	60.9	802	5	AAO20512	Aao20512 Protein o
32	3307	51.2	949	3	AAY44404	Aay44404 Human tru
33	3307	51.2	949	5	AAU79539	Aau79539 Truncated
34	2913	45.1	584	6	ABG73734	Abg73734 Tankyrase
35	2722	42.1	522	4	AAB66287	Aab66287 Human tan
36	2165	33.5	415	4	ABB11320	Abb11320 Human Grb
37	2132.5	33.0	673	3	AAY44403	Aay44403 Human tru
38	2132.5	33.0	673	5	AAU79538	Aau79538 Truncated
39	2056.5	31.8	1099	4	AAB66301	Aab66301 Human tan
40	1967	30.4	373	6	ABU70878	Abu70878 Human adi
41	1942	30.0	385	4	AAB48571	Aab48571 Human T c
42	1895	29.3	362	4	AAB66300	Aab66300 Human TRF
43	1736	26.9	338	6	ABG73733	Abg73733 Tankyrase
44	1706	26.4	331	4	AAB48572	Aab48572 Human bre
45	1467	22.7	308	6	ABU70889	Abu70889 Human adi

ALIGNMENTS

RESULT 1
AAY97749
ID AAY97749 standard; protein; 1240 AA.
XX
AC AAY97749;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 2 protein sequence.
XX
KW Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR N-PSDB; AAA91488.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 22; Fig 4; 63pp; English.
XX
CC This sequence is the Tankyrase homologue isotype 2 (TaHo-2) protein of
CC the invention. The invention also relates to the TaHo-2 protein. The TaHo
CC proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as

CC vaccine, and for construction of hybridisation probes for mapping the
CC gene which encodes TaHo and for the genetic analysis of individuals with
CC genetic disorders. The TaHo proteins, and their coding sequences are
CC useful in screening assays
XX
SQ Sequence 1240 AA;

Query Match 100.0%; Score 6464; DB 4; Length 1240;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240

Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420

Qy        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080

Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||
Db       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
```

RESULT 2
ABG73731
ID ABG73731 standard; protein; 1240 AA.
XX
AC ABG73731;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 2 TaHo-2 fragment from clone K23.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-2;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS Unidentified.
XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
DR N-PSDB; ABQ77067.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Claim 14; Fig 4; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence represents the tankyrase H isoform 2, TaHo-2, isolated from
CC clone K23 which is described in the disclosure of the invention
XX

```
SQ   Sequence 1240 AA;

  Query Match           100.0%;  Score 6464;  DB 6;  Length 1240;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1240;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240

Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420

Qy        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080

Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||
Db       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240


RESULT 3
AAB66290
ID   AAB66290 standard; protein; 1262 AA.
XX
AC   AAB66290;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone consensus protein SEQ ID NO: 107.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63930.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 2; Page 173-176; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1262 AA;

  Query Match           98.6%;  Score 6375.5;  DB 4;  Length 1262;
  Best Local Similarity  99.2%;  Pred. No. 0;
  Matches 1225;  Conservative    1;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          6 RGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 65
              ||: |    ||||  |  ||||||||||||||||||||||||||||||||||||||||||
Db         29 RGSRGAGSPARGAR-GRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 87
```

```
Qy        66 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        88 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 147

Qy       126 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       148 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 207

Qy       186 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       208 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 267

Qy       246 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       268 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 327

Qy       306 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       328 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 387

Qy       366 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       388 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 447

Qy       426 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       448 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 507

Qy       486 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       508 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 567

Qy       546 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       568 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 627

Qy       606 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       628 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 687

Qy       666 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       688 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 747

Qy       726 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       748 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 807

Qy       786 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       808 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 867

Qy       846 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       868 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 927

Qy       906 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       928 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 987

Qy       966 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       988 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1047

Qy      1026 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1048 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1107

Qy      1086 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1108 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1167

Qy      1146 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1168 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1227

Qy      1206 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
             ||||||||||||||||||||||||||||||||||||
Db      1228 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1262
```

RESULT 4
AAB66294
ID AAB66294 standard; protein; 1385 AA.
XX
AC AAB66294;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-LONG SEQ ID NO: 133.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63952.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Claim 2; Page 191-194; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1385 AA;

Query Match 98.6%; Score 6375.5; DB 4; Length 1385;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1225; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```
Qy          6 RGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 65
              ||: |    |||| |  ||||||||||||||||||||||||||||||||||||||||||
Db        152 RGSRGAGSPARGAR-GRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 210

Qy         66 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 270

Qy        126 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 330

Qy        186 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        331 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 390

Qy        246 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 305

Db        391 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 450

Qy        306 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 365

Db        451 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 510

Qy        366 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 425

Db        511 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 570

Qy        426 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 485

Db        571 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 630

Qy        486 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 545

Db        631 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 690

Qy        546 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 605

Db        691 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 750

Qy        606 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 665

Db        751 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 810

Qy        666 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 725

Db        811 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 870

Qy        726 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 785

Db        871 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 930

Qy        786 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 845

Db        931 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 990

Qy        846 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 905

Db        991 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 1050

Qy        906 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 965

Db       1051 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 1110

Qy        966 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1025

Db       1111 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1170

Qy       1026 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1085

Db       1171 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1230

Qy       1086 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1145

Db       1231 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1290

Qy       1146 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1205

Db       1291 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1350

Qy       1206 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240

Db       1351 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1385
```

RESULT 5
AAB66288

```
ID   AAB66288 standard; protein; 1169 AA.
XX
AC   AAB66288;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63926.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 1; Page 162-1665; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1169 AA;
```

Query Match 94.7%; Score 6121; DB 4; Length 1169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         72 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 131

Db          1 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 60

Qy        132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191

Db         61 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 120

Qy        192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 251

Db        121 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 180

Qy        252 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 311

Db        181 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 240

Qy        312 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 371

Db        241 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 300

Qy        372 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 431

Db        301 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 360
```

```
Qy        432 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 420

Qy        492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 480

Qy        552 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 540

Qy        612 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 600

Qy        672 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 660

Qy        732 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 720

Qy        792 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 780

Qy        852 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 840

Qy        912 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 900

Qy        972 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 960

Qy       1032 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1020

Qy       1092 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1080

Qy       1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1140

Qy       1212 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||||||
Db       1141 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1169
```

```
RESULT 6
AAB66278
ID   AAB66278 standard; protein; 1169 AA.
XX
AC   AAB66278;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 related protein sequence SEQ ID NO: 2.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63837.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Disclosure; Page 109-113; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1169 AA;
```

Query Match 94.7%; Score 6121; DB 4; Length 1169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         72 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 60

Qy        132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 120

Qy        192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 180

Qy        252 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 240

Qy        312 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 300

Qy        372 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 360

Qy        432 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 420

Qy        492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 480

Qy        552 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 540

Qy        612 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 600

Qy        672 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 660

Qy        732 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 720

Qy        792 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 780

Qy        852 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 840

Qy        912 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 900

Qy        972 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 960

Qy       1032 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1020

Qy       1092 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1080

Qy       1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1140

Qy       1212 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||||||
Db       1141 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1169
```

RESULT 7
AAB66295
ID AAB66295 standard; protein; 1166 AA.
XX
AC AAB66295;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63953.
XX

PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Claim 3; Page 200-203; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1166 AA;

Query Match 94.5%; Score 6108; DB 4; Length 1166;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy         135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy         195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy         255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy         315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy         375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy         435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy         495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy         555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy         615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy         675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy         735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy         795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780
```

```
Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

RESULT 8
AAB27211
ID AAB27211 standard; protein; 1166 AA.
XX
AC AAB27211;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase II protein sequence SEQ ID NO: 6.
XX
KW Human; tankyrase II; telomere length; signal transduction.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Misc-difference 124
FT   /note= "encoded by TTA"
FT Misc-difference 125
FT   /note= "encoded by TAC"
XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009558.
XX
PR 09-APR-1999; 99US-0128577P.
PR 13-APR-1999; 99US-0129123P.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
DR N-PSDB; AAC66825.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 4; Fig 4; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase II protein
CC sequence. The protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell
XX
SQ Sequence 1166 AA;

Query Match 94.2%; Score 6092; DB 3; Length 1166;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1164; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLLRHGADPNAR 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
              ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DNWNTYPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy        735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900
```

```
Qy       975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
Db       901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy      1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
Db       961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy      1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
Db      1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy      1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
Db      1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy      1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db      1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

RESULT 9
ABP69457
ID ABP69457 standard; protein; 1166 AA.
XX
AC ABP69457;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polypeptide SEQ ID NO 1504.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; ABZ11674.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1504; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1166 AA;

```
  Query Match            94.2%;  Score 6088;  DB 5;  Length 1166;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 1163;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACPNGDVERVKRLVTPEKVNSRDTAGRKIH 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
Db        301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
Db        421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
Db        481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
Db        541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
Db        601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy        735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
Db        661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
Db        721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
```

```
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

```
RESULT 10
AAY72589
ID   AAY72589 standard; protein; 1166 AA.
XX
AC   AAY72589;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human tankyrase homolog protein (THP).
XX
KW   Human; tankyrase homolog protein; THP; gene therapy; cancer; tumour;
KW   basal cell carcinoma; therapy; genetic mapping; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 368
FT                   /label= Met, Leu, Val
FT   Misc-difference 392
FT                   /label= Asn, Thr
FT   Misc-difference 415
FT                   /label= Val, Ile
FT   Misc-difference 558
FT                   /label= Gly, Glu
FT   Misc-difference 559
FT                   /label= Gly, Ala
FT   Misc-difference 764
FT                   /label= Leu, Val
FT   Misc-difference 884
FT                   /label= Asn, His, Asp, Tyr
XX
PN   WO200104326-A1.
XX
PD   18-JAN-2001.
XX
PF   03-JUL-2000; 2000WO-EP006609.
XX
PR   09-JUL-1999;   99US-00350982.
XX
PA   (PHAA ) PHARMACIA & UPJOHN SPA.
XX
PI   Berthelsen J,  Toma S,  Isacchi A;
XX
DR   WPI; 2001-168422/17.
DR   N-PSDB; AAD02578.
XX
PT   New tankyrase homolog protein (THP) polynucleotide and polypeptide useful
PT   in gene therapy, diagnosis and treatment or prevention of unregulated
PT   cell growth, such as cancer or tumor cell growth.
XX
PS   Claim 20; Page 50-54; 60pp; English.
XX
CC   The present sequence is human tankyrase homolog protein (THP). The THP
CC   polypeptides and polynucleotides of the invention are useful in gene
CC   therapy and for treating or preventing unregulated cell growth such as
CC   cancer or tumour (e.g. basal cell carcinoma). The nucleic acid molecules
CC   of the invention and their fragments are useful for restriction fragment
CC   length polymorphism (RFLP) associated with certain disorders, as well as
CC   for genetic mapping. Antisense oligonucleotides, or fragments of nucleic
CC   acid encoding THP are useful as diagnostic tools for probing the
CC   expression of Thp gene in various tissues. THP can be used as antigens
CC   for raising antibodies against them and in assays for identifying
CC   compounds that modulate their activity. They are used in the manufacture
CC   of a medicament directed towards cancers or tumours. THP are also useful
CC   for screening compounds in a variety of drugs screening techniques and as
CC   a research tool for identification, characterisation and purification of
CC   interacting, regulatory proteins
XX
SQ   Sequence 1166 AA;
```

```
  Query Match             93.6%;  Score 6053;  DB 4;  Length 1166;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1157;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy         135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
               ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PLHLAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy         195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy         255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy         315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy         375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy         435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
               ||||||| |||||||||||||||||||||| |||||||||||||||||||| |||||
Db         361 PQTHETAXHCAAASPYPKRKQICELLLRKGAXINEKTKEFLTPLHVASEKAHNDXVEVVV 420

Qy         495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy         555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy         615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
               |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db         541 VVEYLLQHGADVHAKDKXXLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy         675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
               |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db         601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKSCLAR 660

Qy         735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720
```

```
Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALXLAHGADPTLKNQEGQT 780

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRXLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

RESULT 11
AAY97748
ID AAY97748 standard; protein; 1100 AA.
XX
AC AAY97748;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 1 protein sequence.
XX
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR N-PSDB; AAA91487.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 22; Fig 3; 63pp; English.
XX
CC This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein of
CC the invention. The invention also relates to the TaHo-2 protein. The TaHo
CC proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as
CC vaccine, and for construction of hybridisation probes for mapping the
CC gene which encodes TaHo and for the genetic analysis of individuals with
CC genetic disorders. The TaHo proteins, and their coding sequences are
CC useful in screening assays
XX
SQ Sequence 1100 AA;

Query Match 89.2%; Score 5766; DB 4; Length 1100;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1221 AYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100
```

RESULT 12
ABG73732
ID ABG73732 standard; protein; 1100 AA.
XX
AC ABG73732;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 1 TH-1.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TH-1; TaHo;
KW cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS Unidentified.
XX

	Key	Location/Qualifiers
FT	Region	1. .22
FT		/label= Ankyrin_repeat_1
FT	Region	24. .55
FT		/label= Ankyrin_repeat_2
FT	Region	57. .87
FT		/label= Ankyrin_repeat_3
FT	Region	89. .113
FT		/label= Ankyrin_repeat_4
FT	Region	115. .139
FT		/label= Ankyrin_repeat_5
FT	Misc-difference	127
FT		/note= "TH1 start"
FT	Region	147. .176
FT		/label= Ankyrin_repeat_6
FT	Region	178. .208
FT		/label= Ankyrin_repeat_7
FT	Region	210. .241
FT		/label= Ankyrin_repeat_8
FT	Region	243. .281
FT		/label= Ankyrin_repeat_9
FT	Region	283. .332
FT		/label= Ankyrin_repeat_10
FT	Region	334. .365
FT		/label= Ankyrin_repeat_11
FT	Region	367. .398
FT		/label= Ankyrin_repeat_12
FT	Region	400. .424
FT		/label= Ankyrin_repeat_13
FT	Region	426. .458
FT		/label= Ankyrin_repeat_14
FT	Region	460. .491
FT		/label= Ankyrin_repeat_15
FT	Region	493. .524
FT		/label= Ankyrin_repeat_16
FT	Region	526. .557
FT		/label= Ankyrin_repeat_17
FT	Region	559. .594
FT		/label= Ankyrin_repeat_18
FT	Region	596. .644
FT		/label= Ankyrin_repeat_19
FT	Region	646. .677
FT		/label= Ankyrin_repeat_20
FT	Region	679. .710
FT		/label= Ankyrin_repeat_21
FT	Region	712. .755
FT		/label= Ankyrin_repeat_22
FT	Region	757. .793
FT		/label= Ankyrin_repeat_23
FT	Domain	797. .845
FT		/label= SAM_domain
FT	Region	855. .869
FT		/note= "Unidentified region"
FT	Domain	956. .1094
FT		/label= PARP_domain

XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Disclosure; Fig 16A-C; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence represents tankyrase H isoform 1, TH-1, described in the
CC disclosure of the invention
XX
SQ Sequence 1100 AA;

Query Match 89.2%; Score 5766; DB 6; Length 1100;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
```

```
Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1221 AYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||
Db        1081 AYPEYLITYQIMRPEGMVDG 1100
```

```
RESULT 13
ABG73730
ID   ABG73730 standard; protein; 1100 AA.
XX
AC   ABG73730;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TaHo-1 from clone TH-1.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-1;
KW   TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS   Unidentified.
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
DR   N-PSDB; ABQ77066.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Claim 14; Fig 3; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform 1, TaHo-1, isolated from
CC   clone TH-1 which is described in the disclosure of the invention
XX
SQ   Sequence 1100 AA;

  Query Match             89.2%;  Score 5766;  DB 6;  Length 1100;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1221 AYPEYLITYQIMRPEGMVDG 1240
                  ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100


RESULT 14
ABG73735
ID   ABG73735 standard; protein; 1100 AA.
XX
AC   ABG73735;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TH-1 variant Taho F/L.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; mutant;
KW   TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP; variant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .22
FT                   /label= Ankyrin_repeat_1
FT   Region          24. .55
FT                   /label= Ankyrin_repeat_2
FT   Region          57. .87
FT                   /label= Ankyrin_repeat_3
FT   Region          89. .113
FT                   /label= Ankyrin_repeat_4
FT   Region          115. .139
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 127
FT                   /note= "TH1 start"
FT   Region          147. .176
FT                   /label= Ankyrin_repeat_6
FT   Region          178. .208
FT                   /label= Ankyrin_repeat_7
FT   Region          210. .241
FT                   /label= Ankyrin_repeat_8
FT   Region          243. .281
FT                   /label= Ankyrin_repeat_9
FT   Region          283. .332
FT                   /label= Ankyrin_repeat_10
FT   Region          334. .365
FT                   /label= Ankyrin_repeat_11
FT   Region          367. .398
FT                   /label= Ankyrin_repeat_12
FT   Region          400. .424
FT                   /label= Ankyrin_repeat_13
FT   Region          426. .458
FT                   /label= Ankyrin_repeat_14
FT   Region          460. .491
FT                   /label= Ankyrin_repeat_15
FT   Region          493. .524
FT                   /label= Ankyrin_repeat_16
FT   Region          526. .557
FT                   /label= Ankyrin_repeat_17
FT   Region          559. .594
FT                   /label= Ankyrin_repeat_18
FT   Region          596. .644
FT                   /label= Ankyrin_repeat_19
FT   Region          646. .677
FT                   /label= Ankyrin_repeat_20
FT   Region          679. .710
FT                   /label= Ankyrin_repeat_21
FT   Region          712. .755
FT                   /label= Ankyrin_repeat_22
FT   Region          757. .793
FT                   /label= Ankyrin_repeat_23
FT   Domain          797. .845
FT                   /label= SAM_domain
FT   Region          855. .869
FT                   /note= "Unidentified region"
FT   Domain          956. .1094
FT                   /label= PARP_domain
FT   Misc-difference 995
FT                   /note= "This residue is mutated from Phe (wild type, see
FT                   ABG73732) to Leu"
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-C; 90pp; English.
XX
```

```
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform 1, TH-1 variant Taho F/L
CC   described in the disclosure of the invention
XX
SQ   Sequence 1100 AA;

  Query Match             89.1%;  Score 5760;  DB 6;  Length 1100;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1098;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYLAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1221 AYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||
Db        1081 AYPEYLITYQIMRPEGMVDG 1100


RESULT 15
ABG73736
ID   ABG73736 standard; protein; 1084 AA.
XX
AC   ABG73736;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TH-1 variant Taho E/A dC.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TH-1; TaHo;
KW   cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .22
FT                   /label= Ankyrin_repeat_1
FT   Region          24. .55
FT                   /label= Ankyrin_repeat_2
FT   Region          57. .87
FT                   /label= Ankyrin_repeat_3
FT   Region          89. .113
FT                   /label= Ankyrin_repeat_4
FT   Region          115. .139
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 127
FT                   /note= "TH1 start"
FT   Region          147. .176
FT                   /label= Ankyrin_repeat_6
FT   Region          178. .208
FT                   /label= Ankyrin_repeat_7
FT   Region          210. .241
FT                   /label= Ankyrin_repeat_8
FT   Region          243. .281
FT                   /label= Ankyrin_repeat_9
FT   Region          283. .332
FT                   /label= Ankyrin_repeat_10
FT   Region          334. .365
FT                   /label= Ankyrin_repeat_11
FT   Region          367. .398
FT                   /label= Ankyrin_repeat_12
FT   Region          400. .424
FT                   /label= Ankyrin_repeat_13
FT   Region          426. .458
FT                   /label= Ankyrin_repeat_14
FT   Region          460. .491
FT                   /label= Ankyrin_repeat_15
```

```
FT   Region          493. .524
FT                   /label= Ankyrin_repeat_16
FT   Region          526. .557
FT                   /label= Ankyrin_repeat_17
FT   Region          559. .594
FT                   /label= Ankyrin_repeat_18
FT   Region          596. .644
FT                   /label= Ankyrin_repeat_19
FT   Region          646. .677
FT                   /label= Ankyrin_repeat_20
FT   Region          679. .710
FT                   /label= Ankyrin_repeat_21
FT   Region          712. .755
FT                   /label= Ankyrin_repeat_22
FT   Region          757. .793
FT                   /label= Ankyrin_repeat_23
FT   Domain          797. .845
FT                   /label= SAM_domain
FT   Region          855. .869
FT                   /note= "Unidentified region"
FT   Domain          956. .1084
FT                   /label= PARP_domain
FT   Misc-difference 1072
FT                   /note= "This residue is mutated from Glu (wild type, see
FT                   ABG73732) to Ala"
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-C; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform 1, Taho E/A dC, which
CC   contains a Glu to Ala mutation and has a truncated C-terminal containing
CC   a 2 amino acid insertion described in the disclosure of the invention
XX
SQ   Sequence 1084 AA;

  Query Match             87.5%;  Score 5658;  DB 6;  Length 1084;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1079;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAAYVIYRGEQ 1080

Qy       1221 AYPE 1224
              |  :
Db       1081 ALSZ 1084


RESULT 16
AAY05734
ID   AAY05734 standard; protein; 1074 AA.
```

XX
AC AAY05734;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human Grb7 effector 2.2412 protein.
XX
KW Grb7 effector; 2.2412 protein; human; signal transduction; tumour marker;
KW breast cancer; prostate cancer; prognosis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO9915647-A1.
XX
PD 01-APR-1999.
XX
PF 23-SEP-1998; 98WO-AU000795.
XX
PR 23-SEP-1997; 97AU-00009388.
XX
PA (GARV-) GARVAN INST MEDICAL RES.
XX
PI Daly RJ, Sutherland RL;
XX
DR WPI; 1999-254707/21.
DR N-PSDB; AAX25366.
XX
PT New candidate effector for the Grb7 family of signaling proteins, and
PT specific antibody, useful for detection and treatment of cancer.
XX
PS Claim 9; Fig 1; 24pp; English.
XX
CC The present sequence represents a novel candidate effector for the Grb7
CC family of signalling proteins, termed 2.2412. The sequence is predicted
CC from a partial cDNA (see AAX25366). 2.2412 mRNA transcripts (7 kb) were
CC detected in all tissues examined with the exception of kidney. Expression
CC was particularly high in skeletal muscle and placenta. Analysis of the
CC sequence revealed significant homology to a large number of proteins
CC containing ankyrin-like repeats. The 2.2412 gene was localised to between
CC chromosome 10q23.2 and proximal 10q23.32. Deletions in the 10q22-25
CC region have been detected in human breast, prostate, renal, small cell
CC lung and endometrial carcinomas, glioblastoma multiforme, melanoma and
CC meningiomas. Detection of the protein encoded by the 2.2412 cDNA in a
CC sample should provide a useful tumour marker and/or prognostic indicator
CC for certain human cancers, in particular breast cancer and prostate
CC cancer. Antagonism of the interaction between Grb7 family members and the
CC encoded protein should provide a novel treatment stategy for human
CC diseases exhibiting aberrant receptor tyrosine kinase signalling, such as
CC cancer. Anti-Grb7 antibodies can be used in methods of detecting the
CC presence of 2.2412 protein in a sample
XX
SQ Sequence 1074 AA;

Query Match 86.7%; Score 5605; DB 2; Length 1074;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1069; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy        167 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 60

Qy        227 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 120

Qy        287 TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNA 346
              |||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db        121 TPLHLAAGYNRVKIVQLLLQHGRDVHAKDKGDLVPLHNACSYGHYEVTELLVKHGGCVNA 180

Qy        347 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKG 406
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        181 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCKNKSAIDLAPTPQLKERLAYEFKG 240

Qy        407 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 300

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 360

Qy        527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 420

Qy        587 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480

Qy        647 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 540

Qy        707 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 600

Qy        767 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEA 826
              |||||||||||||||||||||||||||||||||||||||| :||||||||||||||||||
Db        601 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFTPLHEA 660

Qy        827 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 720

Qy        887 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 780

Qy        947 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 840

Qy       1007 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 900

Qy       1067 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 960

Qy       1127 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1020

Qy       1187 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1074
```

RESULT 17
AAB47022
ID AAB47022 standard; protein; 1166 AA.
XX
AC AAB47022;
XX
DT 29-MAR-2001 (first entry)
XX
DE Human SPANK.
XX
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW cardiac insufficiency; coronary insufficiency; stroke;

```
KW   high blood pressure; non-insulin dependent diabetes; hypertension;
KW   hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 173. .209
FT                   /note= "Encoded by nucleotides 768. .878"
FT   Misc-difference 327. .362
FT                   /note= "Encoded by nucleotides 1230. .1337"
FT   Misc-difference 375. .398
FT                   /note= "Encoded by nucleotides 1374. .1445"
FT   Misc-difference 482. .524
FT                   /note= "Encoded by nucleotides 1695. .1823"
FT   Misc-difference 641. .677
FT                   /note= "Encoded by nucleotides 2172. .2282"
XX
PN   WO200077225-A1.
XX
PD   21-DEC-2000.
XX
PF   09-JUN-2000; 2000WO-US015926.
XX
PR   11-JUN-1999;   99US-0138957P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Chi N,  Lodish HF;
XX
DR   WPI; 2001-091404/10.
DR   N-PSDB; AAC85294.
XX
PT   New insulin signaling protein SPANK, useful for reducing body mass,
PT   glucose intolerance or insulin resistance and for preventing or treating
PT   obesity-related and muscle-related diseases.
XX
PS   Claim 3; Fig 3; 65pp; English.
XX
CC   This sequence represents human SPANK. The SPANK protein comprises 3
CC   domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP (poly
CC   adenosine diphosphate-ribose polymerase) catalytic domain; and (c) an ANK
CC   domain composed of ankyrin repeats. SPANK is a cytosolic protein which
CC   can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC   aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC   perinuclear region of adipocytes. It is an effector in the insulin
CC   signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC   mass, reducing glucose intolerance or insulin resistance, for preventing
CC   or treating obesity-related diseases or disorders, such as obesity,
CC   cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC   atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC   dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC   also useful for preventing or treating muscle-related diseases or
CC   disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC   Antibodies immunospecific for SPANK are useful for detecting the presence
CC   of SPANK polypeptide in a biological sample
XX
SQ   Sequence 1166 AA;

  Query Match             85.8%;  Score 5543;  DB 4;  Length 1166;
  Best Local Similarity   88.1%;  Pred. No. 0;
  Matches 1092;  Conservative    0;  Mismatches    0;  Indels  148;  Gaps   10;

Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVL-------- 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLKMMALLTP 180

Qy        247 ----------TGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYN 296
                         ||||||||||||||||||||                ||||||||||||
Db        181 LNVNCHASDGTGEYKKDELLESARSGNEE----------------RKSTPLHLAAGYN 222

Qy        297 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH 282

Qy        357 EAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERL---------------- 400
              ||||||||||||||||||||||||||||||||||||||||||||
Db        283 EAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLIKKHLSLEMVNFKHPQ 342

Qy        401 AYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELL 460
              ||||||||||||||||||||              ||||||||||||   |||||||||
Db        343 AYEFKGHSLLQAAREADVTR---------------THETALHCAAAS---KRKQICELL 383

Qy        461 LRKGANINEKTK---EFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 517
              ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        384 LRKGANINEKTKPYPEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 443

Qy        518 GHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLL-----------------QEGI 559
              ||||||||||||||||||||||||||||||||||||||                 ||||
Db        444 GHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLTVKKLCTVQSVNCRDIEGQEGI 503

Qy        560 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 619
              ||||||||||||||||||||                ||||||||||||||||||||
Db        504 SLGNSEADRQLLEAAKAGDVE-----------------RQSTPLHFAAGYNRVSVVEYL 545

Qy        620 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 605

Qy        680 EICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLS 739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 EICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQ-----------------RVKKLS 646

Qy        740 SPDNVNCRDTQG-----------------RHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
              ||||||||||||                ||||||||||||||||||||||||||||||
Db        647 SPDNVNCRDTQGDLLRGDAALLDAAKKGCLARHSTPLHLAAGYNNLEVAEYLLQHGADVN 706

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 766

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 826

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 886

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 946

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW 1006

Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1066

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1126

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||
Db       1127 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

```
RESULT 18
AAB27210
ID   AAB27210 standard; protein; 1431 AA.
XX
AC   AAB27210;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II protein sequence SEQ ID NO: 4.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 41
FT                   /label= Xaa
FT                   /note= "encoded by ANG"
FT   Misc-difference 97
FT                   /label= Xaa
FT                   /note= "encoded by ANG"
FT   Misc-difference 109
FT                   /label= Xaa
FT                   /note= "encoded by NCG"
FT   Misc-difference 113
FT                   /label= Xaa
FT                   /note= "encoded by ANC"
FT   Domain          122. .935
FT                   /note= "ANK domain"
FT   Misc-difference 125
FT                   /label= Xaa
FT                   /note= "encoded by AAN"
FT   Misc-difference 138
FT                   /label= Xaa
FT                   /note= "encoded by NTC"
FT   Misc-difference 165
FT                   /label= Xaa
FT                   /note= "encoded by NGT"
FT   Misc-difference 172. .173
FT                   /label= XaaXaa
FT                   /note= "encoded by NTTRAA"
FT   Misc-difference 177
FT                   /label= Xaa
FT                   /note= "encoded by CAN"
FT   Misc-difference 184
FT                   /label= Xaa
FT                   /note= "encoded by NCA"
FT   Misc-difference 187
FT                   /label= Xaa
FT                   /note= "encoded by NAT"
FT   Misc-difference 201
FT                   /label= Xaa
FT                   /note= "encoded by CMT"
FT   Misc-difference 203. .204
FT                   /label= XaaXaa
FT                   /note= "encoded by RAAANT"
FT   Misc-difference 206
FT                   /label= Xaa
FT                   /note= "encoded by NAT"
FT   Misc-difference 210
FT                   /label= Xaa
FT                   /note= "encoded by CNA"
FT   Misc-difference 212
FT                   /label= Xaa
FT                   /note= "encoded by NGT"
FT   Misc-difference 214
FT                   /label= Xaa
FT                   /note= "encoded by NAM"
FT   Misc-difference 226. .227
FT                   /label= XaaXaa
FT                   /note= "encoded by CNCNAT"
FT   Misc-difference 236. .237
FT                   /label= XaaXaa
FT                   /note= "encoded by GANNNT"
FT   Misc-difference 666
FT                   /label= Xaa
FT                   /note= "encoded by TNT"
FT   Misc-difference 724
FT                   /label= Xaa
FT                   /note= "encoded by AAN"
FT   Misc-difference 726
FT                   /label= Xaa
FT                   /note= "encoded by CNT"
FT   Misc-difference 734
FT                   /label= Xaa
FT                   /note= "encoded by GAN"
FT   Misc-difference 740
FT                   /label= Xaa
FT                   /note= "encoded by NTG"
FT   Misc-difference 762. .763
FT                   /label= XaaXaa
FT                   /note= "encoded by TTNTNT"
FT   Domain          947. .993
FT                   /note= "SAM domain"
FT   Misc-difference 992
FT                   /label= Xaa
FT                   /note= "encoded by NAG"
FT   Misc-difference 1267
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1274
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1277
FT                   /label= Xaa
FT                   /note= "encoded by TGA"
FT   Misc-difference 1279
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1294
FT                   /label= Xaa
FT                   /note= "encoded by TGA"
FT   Misc-difference 1314
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1322. .1323
FT                   /label= XaaXaa
FT                   /note= "encoded by TAATAA"
FT   Misc-difference 1353
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1358
FT                   /label= Xaa
FT                   /note= "encoded by TGA"
FT   Misc-difference 1371
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1377
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1380
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1393
FT                   /label= Xaa
FT                   /note= "encoded by TCA"
FT   Misc-difference 1399
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1403
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1408
FT                   /label= Xaa
```

```
FT                   /note= "encoded by TAG"
FT   Misc-difference 1419
FT                   /label= Xaa
FT                   /note= "encoded by TAA; Xaa=unknown"
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   N-PSDB; AAC66824.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Disclosure; Fig 3; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II protein
CC   sequence. The protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 1431 AA;
```

Query Match 84.7%; Score 5474; DB 3; Length 1431;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 1076; Conservative 6; Mismatches 157; Indels 4; Gaps 2;

```
Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
Db         25 RCLRRRGAAGGQGAHRXGARGRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 84

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
Db         85 RLLALLLAVAAAXIMSGRRCAGGGXACAXAAAEAVEPAARXLFEACRNGDVERXKKLVTP 144

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
Db        145 EKVNSRDTAGRKSTPLHFPAXFGRKDLXXYLLXNGANVQXRDXGGLIPLHNACSFGXAXX 204

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
Db        205 IXLLLXHXAXPNARDNWNYTPXXEAAIKGKIXXCIVLLQHGAEPTIRNTDGRTALDLADP 264

Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
Db        265 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 324

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360
Db        325 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 384

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420
Db        385 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 444

Qy        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
Db        445 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 504

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
Db        505 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 564

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db        565 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 624

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db        625 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVN 684

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
Db        685 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKKKKGXIXLDLVKDGXTDIQDXLRGD 744

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
Db        745 AVLLDAAKKGCLARVKKXXFPDNVNCRDTQGRHSTPLHL-AGXXXXXXXXXXXXXXXXXX 803

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
Db        804 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 863

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
Db        864 XXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALS 923

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
Db        924 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 983

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
Db        984 LEHLMDIFXREQITLDVLVEMGHKELKEIGINAYGHRHKLIKSFERLISGQQGLNPYLTL 1043

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK--- 1077
Db       1044 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNRAKI 1103

Qy       1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
Db       1104 RHEERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1163

Qy       1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
Db       1164 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1223

Qy       1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db       1224 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1266
```

```
RESULT 19
AAE35352
ID   AAE35352 standard; protein; 1083 AA.
XX
AC   AAE35352;
XX
DT   17-JUN-2003  (first entry)
XX
DE   Human colon cancer-associated polypeptide, TNKL (BC-203).
XX
KW   Human; colon cancer-associated polypeptide; immune response; therapy;
KW   colon cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 18
FT                   /note= "Encoded by ATCC"
FT   Misc-difference 77
```

```
FT                   /note= "Encoded by GAGC"
XX
PN   WO200290986-A1.
XX
PD   14-NOV-2002.
XX
PF   02-MAY-2002; 2002WO-US013994.
XX
PR   04-MAY-2001; 2001US-00849602.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chen Y,  Old LJ,  Scanlan MJ,  Stockert E;
XX
DR   WPI; 2003-112003/10.
DR   N-PSDB; AAD54031.
XX
PT   Diagnosing colon cancer in a subject comprises identifying colon cancer-
PT   associated polypeptides as antigens that elicit immune response in colon
PT   cancer.
XX
PS   Example 1; Page 118-121; 122pp; English.
XX
CC   The invention relates to a method for diagnosing colon cancer in a
CC   subject which comprises identifying colon cancer-associated polypeptides
CC   as antigens that elicit immune response in colon cancer. The method is
CC   useful for diagnosing, determining onset, progression, or regression of
CC   colon cancer in a subject, or for selecting a course of treatment of a
CC   subject having or suspected of having colon cancer. The colon cancer-
CC   associated polypeptides are useful as markers for diagnosing colon
CC   cancer, and for following the course of treatment of colon cancer. The
CC   present sequence is human colon cancer-associated polypeptide
XX
SQ   Sequence 1083 AA;

  Query Match             82.0%;  Score 5303.5;  DB 6;  Length 1083;
  Best Local Similarity   93.2%;  Pred. No. 0;
  Matches 1032;  Conservative    6;  Mismatches   22;  Indels   47;  Gaps    3;

Qy         12 QGAQRGARVGAA-----------HGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLL 59
              ||||||||||||           |  |: | :|     |  | :|  :|  :||
Db          1 QGAQRGARVGAAMGLRRSGDSREPSGPGPERVFSGGPRPPARGAGAPAPVAGAVAG---- 56

Qy         60 LRLLALLLAVAAARIMSGRRCAGG-----------GAACASAAAEAVEPAARELFEACRN 108
                              | ||            |:    ||||||||||||||||||
Db         57 --------------------CGGGQDHVGSPLRRRGSGLRDAAAEAVEPAARELFEACRN 96

Qy        109 GDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIP 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIP 156

Qy        169 LHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 LHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 216

Qy        229 TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP 276

Qy        289 LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD 336

Qy        349 LWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 LWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS 396

Qy        409 LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN 456

Qy        469 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 516

Qy        529 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 576

Qy        589 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 636

Qy        649 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD 696

Qy        709 GDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 GDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 756

Qy        769 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQ 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQ 816

Qy        829 KGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 KGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG 876

Qy        889 VRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGV 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 VRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGV 936

Qy        949 DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLI 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLI 996

Qy       1009 SGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNI 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 SGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNI 1056

Qy       1069 LKIQKVCNKKLWERYTHRRKEVSEENH 1095
              ||||||||||||||||||||||||||
Db       1057 LKIQKVCNKKLWERYTHRRKEVSEENH 1083


RESULT 20
AAY44402
ID   AAY44402 standard; protein; 1327 AA.
XX
AC   AAY44402;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human tankyrase.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   two-hybrid screen; telomere length regulation; recombinant tankyrase;
KW   aging; skin atrophy; macular degeneration; atherosclerosis;
KW   ataxia telangiectasia; tankyrase modulator; rational drug design.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          181. .1010
FT                   /note= "Ankyrin-specific (ANK) repeat consensus domain"
FT   Domain          1023. .1088
FT                   /label= Sterile_alpha_motif
FT   Domain          1176. .1314
FT                   /note= "Poly(ADP-ribose)polymerase (PARP) related domain"
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
```

```
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   N-PSDB; AAZ29627.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Claim 1; Fig 1B; 133pp; English.
XX
CC   The present sequence is human tankyrase, a TRF1 (telomeric repeat binding
CC   factor-1) binding protein. This was isolated using a two-hybrid screen
CC   with TRF1. Tankyrase is involved in regulation of telomere length through
CC   TRF1 and may directly modulate the effect of TRF1. Recombinant tankyrase,
CC   or its fragments are used to identify specific modulators which are
CC   potential drugs for countering telomere shortening associated with aging
CC   (e.g. atrophy of skin, macular degeneration or atherosclerosis) or
CC   diseases like ataxia telangiectasia. They are also used to raise specific
CC   antibodies (used as immunoassay reagents and as modulators of tankyrase
CC   activity) and in rational drug design
XX
SQ   Sequence 1327 AA;

  Query Match             79.0%;  Score 5103.5;  DB 3;  Length 1327;
  Best Local Similarity   79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELPEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235
Db       1299 QAYPEYLITYQIMKPE 1314
```

```
RESULT 21
AAB27212
ID   AAB27212 standard; protein; 1327 AA.
XX
AC   AAB27212;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase I protein sequence SEQ ID NO: 8.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
```

PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 4; Fig 5; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase I protein
CC sequence. The invention relates to the isoaltion of the protein and
CC coding sequences of human tankyrase II. This protein is thought to be
CC involved in signal transduction in the cell, and to have binding activity
CC for other telomere-associated proteins. It is possible that it plays a
CC role in the regulation of telomere length, thus affecting the replicative
CC ability of the cell. The protein is useful for ribosylating target
CC proteins, for determining tankyrase II binding activity in a sample, and
CC for modulating telomere length in a cell
XX
SQ Sequence 1327 AA;

Query Match 79.0%; Score 5103.5; DB 3; Length 1327;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

```
Qy          22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
               :| | ||:|  :|| : : |:|     :   : ||      : |     | :|
Db         112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy          82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
               | |||       ||  | ||| |||||||| ||||||    ||:: |||||:|||||||
Db         166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy         142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
               ||||||||:||| |||| ||||||||||||||||||||:||| |||||||||||||||
Db         226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy         202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
               |||||||||||||||||||:| |||||||::|||||||||||||||||||||||||:|||
Db         286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy         262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
               |||||:|||||||||||||||||||||||||||||:|||||||||||||||||| |||
Db         346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy         322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
               |||||||||||||||:|||||||||||||||||||||||||||||||||||||||:|||||:|
Db         406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy         382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
               || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db         466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy         442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
               |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db         526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy         502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
               ||| ||||:|||||  |||||||||||| ||:||||||||| ||||| |||:| |   :
Db         586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy         562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
                 |: | :||||:||||:||||:||: |:|||||:||| |||||||||||||||||||
Db         646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy         622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
               |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db         706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy         682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
               |||||:|||||||||||||||||:|||||||:|||||||||||||||||:|| :|
Db         766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy         742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
               :|:|||||||:|||||||||||||||||||||:|||||||||||||||||||||||||:
Db         826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy         802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
               |||||||| ||||||||||||||||||||||||||||||||||:|||||||||| :|
Db         886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy         862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
               ||: ||| |||| |||:|:|||           : :| |  |:|| ||||||:|||:|
Db         946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy         922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
               :||:  :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db         999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy         980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
               :|||:|||||||||||||||||||||: |||| ||||| :    ||||:||:|:|||:
Db        1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy        1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
               |||||||||:|||||:|||||||:::|||| |||| ||: ||:|||||||||| |
Db        1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy        1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
               ||||||||||:|||||||||||||||||||||||||||||||||||||||||||| ||
Db        1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy        1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
               ||||||||||:|||||||||||||| |||||:||||||| |||||||| |||||||||
Db        1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy        1220 QAYPEYLITYQIMRPE 1235
               |||||||||||||:||
Db        1299 QAYPEYLITYQIMKPE 1314
```

RESULT 22
AAB66279
ID AAB66279 standard; protein; 1327 AA.
XX
AC AAB66279;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase1 SEQ ID NO: 4.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63838.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 118-121; 242pp; English.

```
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1327 AA;

  Query Match             79.0%;  Score 5103.5;  DB 4;  Length 1327;
  Best Local Similarity   79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |      | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||       ||  | ||| |||||||| |||||    ||::| |||||:||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||:||| ||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              ||||||||||||||||||:| |||||||::|||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:|||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
              |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||:|||||||:|||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
              :|:|||||||:|||||||||||||||||:|||||||||||||||||||||||:
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
              |||||||| |||||||||||||||||||||||||||||||||:|||||||||| :|
Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
```

```
                ||: |||  |||| |||:|:|||          : :| |  |:|| ||||||:|||:|
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
              :||:   :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
              :|||:|||||||||||||||||||||: |||| ||||| :    ||||:||:|:|||:
Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
              |||||||||:|||||||:|||||||||:::|||| |||| ||: ||:|||||||||| |
Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
              |||||||||:|||||||||||||| |||||:|||||| |||||||| |||||
Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235
              ||||||||||||:||
Db       1299 QAYPEYLITYQIMKPE 1314


RESULT 23
AAU79537
ID   AAU79537 standard; protein; 1327 AA.
XX
AC   AAU79537;
XX
DT   12-AUG-2002  (first entry)
XX
DE   Human tankyrase.
XX
KW   Human; tankyrase; TRF1 interacting ankyrin; cytostatic; ADP ribosylation;
KW   telomere; nuclear localisation signal; NLS;
KW   telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW   drug development; aging; ataxia telangiectasia; enzyme.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .180
FT                   /label= HPS_domain
FT                   /note= "domain rich in Pro, His and Ser"
FT   Domain          181. .1010
FT                   /label= ANK_repeat_domain
FT   Region          336. .1163
FT                   /label= TRIL4
FT                   /note= "Recombinant protein used for antibody production"
FT   Region          436. .797
FT                   /label= TRIL12
FT                   /note= "Recombinant protein used for antibody production"
FT   Region          973. .1149
FT                   /label= ANK2
FT                   /note= "Recombinant protein used for antibody production"
FT   Region          1023. .1088
FT                   /label= SAM_motif
FT   Domain          1176. .1314
FT                   /label= PARP-related_domain
XX
PN   WO200229017-A2.
XX
PD   11-APR-2002.
XX
PF   25-SEP-2001; 2001WO-US029931.
XX
PR   03-OCT-2000; 2000US-00678411.
XX
```

```
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2002-463230/49.
DR   N-PSDB; ABK48629.
XX
PT   Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT   to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT   ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT   -ribosylation activity.
XX
PS   Claim 3; Fig 1B; 151pp; English.
XX
CC   The invention discloses a nuclear localisation signal (NLS)-telomeric
CC   protein tankyrase that comprises a NLS and at least a portion of a
CC   tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC   factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC   Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC   TRF1, which is a negative regulator of telomere length maintenance, and
CC   in so doing, counteracts its effect and play a role in lifecycle of a
CC   cell. The protein is useful for identifying agents that modulate the
CC   binding of TRF1 to its DNA binding site, for identifying agents that
CC   modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC   identifying the agents that modulate telomere length and for identifying
CC   agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC   the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC   would hinder telomere lengthening and so is useful in the treatment of
CC   cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC   telomere lengthening and so could be useful for extending the lifetime of
CC   the cell, especially self-renewing cells such as bone marrow cells,
CC   peripheral blood, T and B cells. Tankyrases are also useful for
CC   developing drugs that can counteract the telomere shortening associated
CC   with aging and certain diseases such as ataxia telangiectasia. The
CC   sequence presented is the human tankyrase
XX
SQ   Sequence 1327 AA;

  Query Match             79.0%;  Score 5103.5;  DB 5;  Length 1327;
  Best Local Similarity   79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |      | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||        ||  | ||| |||||||| ||||||     ||:: | |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||||:||  |||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| |||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||||:|||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||:|||||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:|:|::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:|: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:||||| |||||||||||| ||:|||||||| ||||| |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
              |: | :||||:||||:||||:||: |:|||||:||| |||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              |||||||||||||||||||||||||||||||:||| |||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||:|||||||:|||||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801
              :|:||||||||:|||||||||||||||||||||:|||||||||||||||||||||||||:
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
              |||||||| ||||||||||||||||||||||||||||||||||:|||||||||| :|
Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
              ||: |||  |||| |||:|:|||          : :| |  |:|| ||||||:|||:|
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979
               :||:   :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039
              :|||:|||||||||||||||||||||||: |||| |||| :    ||||:||:|:|||:
Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099
              |||||||||:|||||||:|||||||:::|||| |||| ||: ||:|||||||||| |
Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159
              |||||||||:|||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219
              ||||||||||:|||||||||||||| |||||:||||||| |||||||| |||||||||
Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235
              |||||||||||||:||
Db       1299 QAYPEYLITYQIMKPE 1314


RESULT 24
AAB27209
ID   AAB27209 standard; protein; 1333 AA.
XX
AC   AAB27209;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II SEQ ID NO: 2.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .180
FT                   /note= "ANK domain"
FT   Domain          181. .1017
FT                   /note= "ANK domain"
```

```
FT   Misc-difference 199
FT                   /label= Xaa
FT                   /note= "encoded by CNN"
FT   Misc-difference 208. .209
FT                   /label= XaaXaa
FT                   /note= "encoded by ANCCNC"
FT   Misc-difference 212
FT                   /label= Xaa
FT                   /note= "encoded by GNG"
FT   Misc-difference 232
FT                   /label= Xaa
FT                   /note= "encoded by NTT"
FT   Misc-difference 242
FT                   /label= Xaa
FT                   /note= "encoded by NTC"
FT   Misc-difference 244
FT                   /label= Xaa
FT                   /note= "encoded by NCA"
FT   Misc-difference 255
FT                   /label= Xaa
FT                   /note= "encoded by ANT"
FT   Misc-difference 264
FT                   /label= Xaa
FT                   /note= "encoded by ANT"
FT   Misc-difference 726
FT                   /label= Xaa
FT                   /note= "encoded by TNT"
FT   Misc-difference 794
FT                   /label= Xaa
FT                   /note= "encoded by GAN"
FT   Misc-difference 800
FT                   /label= Xaa
FT                   /note= "encoded by NTG"
FT   Misc-difference 822. .823
FT                   /label= XaaXaa
FT                   /note= "encoded by TTNTNT"
FT   Domain          1030. .1095
FT                   /note= "SAM domain"
FT   Misc-difference 1055
FT                   /label= Xaa
FT                   /note= "encoded by NAG; Xaa=unknown"
FT   Domain          1190. .1328
FT                   /note= "PARP domain"
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   N-PSDB; AAC66823.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Disclosure; Fig 2; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II protein
CC   sequence. The protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 1333 AA;

  Query Match            78.0%;  Score 5040.5;  DB 3;  Length 1333;
  Best Local Similarity  86.8%;  Pred. No. 0;
  Matches  984;  Conservative    8;  Mismatches  133;  Indels    9;  Gaps    2;

Qy       116 RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSF 175
Db       200 RLVTPEKVNXXNTXGRKSTPLHFPAGFGRKNLXKYLLQNGANXQXLYNGGLIPLHXACSF 259

Qy       176 GHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTAL 235
Db       260 GHAKXINLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTAL 319

Qy       236 DLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY 295
Db       320 DLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY 379

Qy       296 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 355
Db       380 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 439

Qy       356 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 415
Db       440 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 499

Qy       416 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 475
Db       500 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 559

Qy       476 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 535
Db       560 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 619

Qy       536 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 595
Db       620 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 679

Qy       596 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 655
Db       680 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKH 739

Qy       656 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD 715
Db       740 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTLLDLVKDGXTDIQD 799

Qy       716 LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA--AGYNNLEVAEYLL 773
Db       800 XLRGDAVLLDAAKKGCLARVKKXXFPDNVNCRDTQGRHSTPLHLAGXXXXXXXXXXXXXX 859

Qy       774 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 833
Db       860 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 919

Qy       834 LCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPG 893
Db       920 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPG 979

Qy       894 ATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSIT 953
Db       980 ATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSIT 1039

Qy       954 QFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 1013
Db      1040 QFVRNLGLEHLMDIFXREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 1099

Qy      1014 LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQK 1073
Db      1100 LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQK 1159

Qy      1074 VCNK-------KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1126
```

```
              |||        :  |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 VCNXXXXXAKIRHEERYTHRRKEVSEENHNHANERMLPHGSPFVNAIIHKGFDERHAYIG 1219

Qy       1127 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1279

Qy       1187 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1280 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1333


RESULT 25
AAB47023
ID   AAB47023 standard; protein; 991 AA.
XX
AC   AAB47023;
XX
DT   29-MAR-2001  (first entry)
XX
DE   Mouse SPANK.
XX
KW   SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW   poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW   ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW   adipocyte; insulin signalling pathway; hyperlipidaemia;
KW   glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW   cardiac insufficiency; coronary insufficiency; stroke;
KW   high blood pressure; non-insulin dependent diabetes; hypertension;
KW   hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 483
FT                   /note= "Encoded by G"
FT   Misc-difference 614. .623
FT                   /note= "Encoded by G"
FT   Misc-difference 776
FT                   /note= "Encoded by ATGCC"
XX
PN   WO200077225-A1.
XX
PD   21-DEC-2000.
XX
PF   09-JUN-2000; 2000WO-US015926.
XX
PR   11-JUN-1999;   99US-0138957P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Chi N,  Lodish HF;
XX
DR   WPI; 2001-091404/10.
DR   N-PSDB; AAC85295.
XX
PT   New insulin signaling protein SPANK, useful for reducing body mass,
PT   glucose intolerance or insulin resistance and for preventing or treating
PT   obesity-related and muscle-related diseases.
XX
PS   Claim 3; Fig 5B; 65pp; English.
XX
CC   This sequence represents the mouse SPANK protein. The SPANK protein
CC   comprises 3 domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP
CC   (poly adenosine diphosphate-ribose polymerase) catalytic domain; and (c)
CC   an ANK domain composed of ankyrin repeats. SPANK is a cytosolic protein
CC   which can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC   aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC   perinuclear region of adipocytes. It is an effector in the insulin
CC   signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC   mass, reducing glucose intolerance or insulin resistance, for preventing
CC   or treating obesity-related diseases or disorders, such as obesity,
CC   cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC   atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC   dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC   also useful for preventing or treating muscle-related diseases or
CC   disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC   Antibodies immunospecific for SPANK are useful for detecting the presence
CC   of SPANK polypeptide in a biological sample
XX
SQ   Sequence 991 AA;

  Query Match             77.9%;  Score 5038.5;  DB 4;  Length 991;
  Best Local Similarity   83.3%;  Pred. No. 0;
  Matches  971;  Conservative   10;  Mismatches   10;  Indels  175;  Gaps    2;

Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
              |||||||||||||||| |||||||:|||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAGAEAVEPSARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
              |||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLQHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
              ||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGDYKKDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLHHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
              |||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRIEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
              |||||||||||||||||| |||||||:|||||||||||||||||||||||||||||||||
Db        301 ADPTLLNCHNKSAIDLAPTAQLKERLSYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
              |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANTNEKTKEFLTPLHVASENAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
              ||||||||||:|||||||||:|||||||||||||||||||||||||||||||||||||||
Db        421 KHEAKVNALDSLGQTSLHRAAHCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
              ||                          |||||||||||||||||||||||||||||||
Db        481 LQG----------------------ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 518

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 578

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
              |||||||||||||||||||||||||||||||||||| |  || ||||||||||||||||
Db        579 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTQIGKLLIGDAALLDAAKKGCLAR 638

Qy        735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 698

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 758

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
              ||||||||||||||||
Db        759 PLDLVSADDVSALLTAAM---------------------------------------- 776

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974

Db        777 ---------------------------------------------------------- 776
```

```
Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
                                                                ||||||||
Db        777 ----------------------------------------------------GTILIDLSP 785

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
              ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db        786 DDKEFQSVEEEMQSTVREHRDGGHAGGVFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 845

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 905

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 CPIHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 965

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||:||||||||
Db        966 IYRGEQAYPEYLITYQIVRPEGMVDG 991
```

RESULT 26
AAB48574
ID AAB48574 standard; protein; 907 AA.
XX
AC AAB48574;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human breast cancer MO-BC-203 polypeptide.
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; MO-BC-203.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR N-PSDB; AAC89706.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 8; Page 84-86; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 907 AA;

Query Match 71.5%; Score 4624.5; DB 4; Length 907;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 885; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

```
Qy        333 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 60

Qy        393 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 120

Qy        453 RKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 512
              |||||||||:||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RKQICELLLKKGSNINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 180

Qy        513 RAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 240

Qy        573 AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 300

Qy        633 GLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADP 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADP 360

Qy        693 TKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGR 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGR 420

Qy        753 HSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 HSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 480

Qy        813 NATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM 872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM 540

Qy        873 PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSS 932
              |||||||||||||||||||||||||     :  :    :|  :   |||||||||||
Db        541 PPSALPSCYKPQVLNGVRSPGATADALFQVHLAHQAFLQPAVLTTI-WEFSELSSVVSSS 599

Qy        933 GTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN 992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN 659

Qy        993 AYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVRE 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVRE 719

Qy       1053 HRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNA 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 HRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNA 779

Qy       1113 IIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLF 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 IIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLF 839

Qy       1173 CRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIM 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIM 899

Qy       1233 RPEGMVDG 1240
              ||||||||
Db        900 RPEGMVDG 907
```

RESULT 27

```
AAB66285
ID   AAB66285 standard; protein; 784 AA.
XX
AC   AAB66285;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone protein sequence SEQ ID NO: 89.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63917.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 1; Page 144-146; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 784 AA;

  Query Match             63.5%;  Score 4102;  DB 4;  Length 784;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  784;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        457 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 60

Qy        517 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 120

Qy        577 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 180

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 240

Qy        697 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 300

Qy        757 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 360

Qy        817 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 420

Qy        877 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 480

Qy        937 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 540

Qy        997 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 600

Qy       1057 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 660

Qy       1117 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 720

Qy       1177 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 780

Qy       1237 MVDG 1240
              ||||
Db        781 MVDG 784


RESULT 28
ABB60894
ID   ABB60894 standard; protein; 1181 AA.
XX
AC   ABB60894;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 9474.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL04997.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 9474; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
```

```
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1181 AA;

  Query Match           62.0%;  Score 4005;  DB 4;  Length 1181;
  Best Local Similarity  65.3%;  Pred. No. 0;
  Matches  772;  Conservative  166;  Mismatches  189;  Indels   56;  Gaps   13;

Qy         91 AAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEY 150
Db         17 AVMANDP-LRELFEACKTGEIAKVKKLITPQTVNARDTAGRKSTPLHFAAGYGRREVVEF 75

Qy        151 LLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGK 210
Db         76 LLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYTPLHEAASKGK 135

Qy        211 IDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMAL 270
Db        136 VDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAARSGAEDRLLAL 195

Qy        271 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGH 330
Db        196 LTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLVPLHNACSYGH 255

Qy        331 YEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL 390
Db        256 FDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLLNCHSKSAIDA 315

Qy        391 APTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPY 450
Db        316 APTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDTPLHLAVVSPD 375

Qy        451 PKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTS 510
Db        376 GKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKVNALDSLGQTP 435

Qy        511 LHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQL 570
Db        436 LHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK-----NPPDSETHL 488

Qy        571 LEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 628
Db        489 LEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQFLLEHGAEVYA 548

Qy        629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 688
Db        549 ADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGKYDICKLLLKH 608

Qy        689 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 748
Db        609 GADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRLVTPESINCRD 668

Qy        749 TQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKY 808
Db        669 AQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGHLDIAALLIKH 728

Qy        809 NACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL 868
Db        729 KTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIELATADDVKCLL 788

Qy        869 TAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAASSLDNLSGSF 922
Db        789 QDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPTTETVLLPTGA 847

Qy        923 SELSSV-----VSSS------------GTEGASSLEKKEVPGVD--FSITQFVRNLGLEH 963
```

```
Db        848 SMILSVPVPLPLSSSTRISPAQGAEANGAEGSSSDDL--LPDADTITNVSGFLSSQQLHH 905

Qy        964 LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTS 1023
Db        906 LIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS-----------TT 953

Qy       1024 GSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK 1077
Db        954 GIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYNIIRVQKVQNR 1013

Qy       1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
Db        1014 KLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGGMFGAGIYFAE 1073

Qy       1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
Db        1074 HSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAMKMAHAPPGHH 1133

Qy       1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db        1134 SVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176
```

```
RESULT 29
AAB66297
ID   AAB66297 standard; protein; 1181 AA.
XX
AC   AAB66297;
XX
DT   05-APR-2001  (first entry)
XX
DE   Drosophila tankyrase SEQ ID NO: 139.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Drosophila melanogaster.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63955.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 2; Page 215-218; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1181 AA;

  Query Match           61.8%;  Score 3997;  DB 4;  Length 1181;
  Best Local Similarity  65.2%;  Pred. No. 0;
  Matches  771;  Conservative  166;  Mismatches  190;  Indels   56;  Gaps   13;
```

```
Qy        91 AAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEY 150
Db        17 AVMANDP-LRELSEACKTGEIAKVKKLITPQTVNARDTAGRKSTPLHFAAGYGRREVVEF 75

Qy       151 LLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGK 210
Db        76 LLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYTPLHEAASKGK 135

Qy       211 IDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMAL 270
Db       136 VDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAARSGAEDRLLAL 195

Qy       271 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGH 330
Db       196 LTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLVPLHNACSYGH 255

Qy       331 YEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL 390
Db       256 FDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLLNCHSKSAIDA 315

Qy       391 APTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPY 450
Db       316 APTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDTPLHLAVVSPD 375

Qy       451 PKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTS 510
Db       376 GKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKVNALDSLGQTP 435

Qy       511 LHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQL 570
Db       436 LHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK-----NPPDSETHL 488

Qy       571 LEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 628
Db       489 LEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQFLLEHGAEVYA 548

Qy       629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 688
Db       549 ADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGKYDICKLLLKH 608

Qy       689 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 748
Db       609 GADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRLVTPESINCRD 668

Qy       749 TQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKY 808
Db       669 AQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGHLDIAALLIKH 728

Qy       809 NACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL 868
Db       729 KTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIELATADDVKCLL 788

Qy       869 TAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAASSLDNLSGSF 922
Db       789 QDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPTTETVLLPTGA 847

Qy       923 SELSSV-----VSSS------------GTEGASSLEKKEVPGVD--FSITQFVRNLGLEH 963
Db       848 SMILSVPVPLPLSSSTRISPAQGAEANGAEGSSSDDL--LPDADTITNVSGFLSSQQLHH 905

Qy       964 LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTS 1023
Db       906 LIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS-----------TT 953

Qy      1024 GSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK 1077
Db       954 GIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYNIIRVQKVQNR 1013

Qy      1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
Db       1014 KLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGGMFGAGIYFAE 1073

Qy      1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
Db       1074 HSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAMKMAHAPPGHH 1133

Qy      1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db       1134 SVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176
```

RESULT 30
AAB66286
ID AAB66286 standard; protein; 756 AA.
XX
AC AAB66286;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone protein sequence SEQ ID NO: 91.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63918.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 150-152; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 756 AA;

```
  Query Match             61.2%;  Score 3959;  DB 4;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  756;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       485 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 544
Db         1 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 60

Qy       545 QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPL 604
Db        61 QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPL 120

Qy       605 HFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL 664
Db        121 HFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL 180

Qy       665 WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALL 724
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALL 240

Qy        725 DAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 784
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 300

Qy        785 GGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGAD 844
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGAD 360

Qy        845 PTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPS 904
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPS 420

Qy        905 SPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 964
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 480

Qy        965 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 1024
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 540

Qy       1025 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 1084
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 600

Qy       1085 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 1144
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 660

Qy       1145 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 1204
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 720

Qy       1205 VNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
                 ||||||||||||||||||||||||||||||||||||
Db        721 VNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 756
```

RESULT 31
AAO20512

```
ID   AAO20512 standard; protein; 802 AA.
XX
AC   AAO20512;
XX
DT   27-JUN-2002  (first entry)
XX
DE   Protein of APP related human homologue hCP33787.
XX
KW   Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW   amyloid precursor protein; tissue-specific expression control; human APP;
KW   APP pathway modulator; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200226820-A2.
XX
PD   04-APR-2002.
XX
PF   01-OCT-2001; 2001WO-EP011345.
XX
PR   29-SEP-2000; 2000US-0236893P.
PR   14-JUN-2001; 2001US-0298309P.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   Cohen D,  Dengler UJ,  Finelli AL,  Freuler F,  Konsolaki M;
PI   Reinhardt MWHM,  Zusman S;
XX
DR   WPI; 2002-315796/35.
DR   N-PSDB; AAK99406.
XX
PT   New transgenic fly, containing DNA encoding an Abeta portion of human
PT   APP, useful for identifying agents which modulate the APP pathway and
PT   which can be used to treat Alzheimer's disease.
XX
PS   Example 4; Page 113-114; 129pp; English.
XX
CC   The invention relates to a transgenic fly whose genome comprises DNA
CC   encoding a polypeptide having the Abeta portion of human amyloid
CC   precursor protein (APP), fused to a signal sequence. The DNA sequence
CC   encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC   the specification. The DNA sequence is operably linked to a tissue-
CC   specific expression control sequence. Expression of the sequence gives
CC   the fly an altered phenotype. The purpose of the invention is for
CC   identifying agents that inhibit or promote the expression and/or function
CC   of genes or encoded polypeptides which modify the APP pathway. The agent
CC   is a compound, triple helix DNA, antisense oligonucleotide, double
CC   stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC   to treat conditions such as Alzheimer's disease. The agent can be used as
CC   an APP pathway modulator or in gene therapy. This sequence represents the
CC   protein of the APP related human homologue hCP33787
XX
SQ   Sequence 802 AA;
```

Query Match 60.9%; Score 3938; DB 5; Length 802;
Best Local Similarity 67.2%; Pred. No. 0;
Matches 787; Conservative 4; Mismatches 5; Indels 376; Gaps 5;

```
Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGE----- 249
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGKSVYSG 180

Qy        250 YKKDELLES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHG 308
                |.:  :  : ||||||||||||||||||||||||||
Db        181 YSRKPVKNNLARSGNEEKMMALLTPLNVNCHASDGR---------------------- 216

Qy        309 ADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCS 368
                                                ||||||||||||||||||||||||||||
Db        217 ----------------------------KHGACVNAMDLWQFTPLHEAASKNRVEVCS 246

Qy        369 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE 428
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE 306

Qy        429 MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHND 488
                ||||||||||||||
Db        307 MVNFKHPQTHETAL---------------------------------------------- 320

Qy        489 VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN 548
Db        321 ------------------------------------------------------------ 320

Qy        549 ENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAA 608
                                                ||||||||||||||||||||||||
Db        321 ---------------------------------KLCTVQSVNCRDIEGRQSTPLHFAA 345

Qy        609 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 668
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 405

Qy        669 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 728
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 465

Qy        729 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 788
```

```
Db        466 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 525

Qy        789 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 848
              |||||||||
Db        526 PLHNAASYG---------------------------------------------------- 534

Qy        849 NQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSS 908

Db        535 ---------------------------------------------------------- 534

Qy        909 LSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIF 968

Db        535 ---------------------------------------------------------- 534

Qy        969 EREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTI 1028
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ----ITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTI 590

Qy       1029 LIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 LIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 650

Qy       1089 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 710

Qy       1149 IGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGL 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 IGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGL 770

Qy       1209 ALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||
Db        771 ALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 802
```

RESULT 32
AAY44404
ID AAY44404 standard; protein; 949 AA.
XX
AC AAY44404;
XX
DT 22-MAR-2000 (first entry)
XX
DE Human truncated tankyrase-2.
XX
KW Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW telomere length regulation; recombinant tankyrase; aging; skin atrophy;
KW macular degeneration; atherosclerosis; ataxia telangiectasia;
KW tankyrase modulator; rational drug design.
XX
OS Homo sapiens.
XX
PN WO9964606-A1.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-US012968.
XX
PR 10-JUN-1998; 98US-00095225.
PR 17-AUG-1998; 98US-00135233.
PR 19-NOV-1998; 98US-00196387.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2000-116549/10.
DR N-PSDB; AAZ29633.
XX
PT New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT length, used to identify modulators, e.g. for inhibiting growth of
PT cancer.
XX
PS Disclosure; Page 125-129; 133pp; English.
XX
CC The present sequence is truncated human tankyrase, a TRF1 (telomeric
CC repeat binding factor-1) binding protein. This comprises 1-881 residues
CC of human tankyrase. Tankyrase is involved in regulation of telomere
CC length through TRF1 and may directly modulate the effect of TRF1.
CC Recombinant tankyrase, or its fragments are used to identify specific
CC modulators which are potential drugs for countering telomere shortening
CC associated with aging (e.g. atrophy of skin, macular degeneration or
CC atherosclerosis) or diseases like ataxia telangiectasia. They are also
CC used to raise specific antibodies (used as immunoassay reagents and as
CC modulators of tankyrase activity) and in rational drug design
XX
SQ Sequence 949 AA;

Query Match 51.2%; Score 3307; DB 3; Length 949;
Best Local Similarity 76.2%; Pred. No. 1.6e-285;
Matches 638; Conservative 72; Mismatches 119; Indels 8; Gaps 2;

```
Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |      |  :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||      ||  | ||| |||||||| ||||||    ||::| ||||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              |||||||||:||| |||| ||||||||||||||||||||||:|||  ||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| |||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||:|||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: ||||||||||:||| |:|:|||||||:|:|||||:|| || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||||| ||:|||||||| ||||| |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
                |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              |||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||:|||||||:|||||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
              :|:|||||||:|||||||||||||||||:|||||||||||||||||||||||||||
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCLA 885

Qy        800 DVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856
```

```
Db        886 RVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942
```

RESULT 33
AAU79539
ID AAU79539 standard; protein; 949 AA.
XX
AC AAU79539;
XX
DT 12-AUG-2002 (first entry)
XX
DE Truncated human tankyrase #2.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; cytostatic; ADP ribosylation;
KW telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia; enzyme.
XX
OS Homo sapiens.
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR N-PSDB; ABK48635.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 146-148; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC protein sequence presented is the truncated human tankyrase #2. Note:
CC Residues 1-881 are the same as the wild-type human tankyrase sequence
CC (AAU79537)
XX
SQ Sequence 949 AA;

Query Match 51.2%; Score 3307; DB 5; Length 949;
Best Local Similarity 76.2%; Pred. No. 1.6e-285;
Matches 638; Conservative 72; Mismatches 119; Indels 8; Gaps 2;

```
Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCLA 885

Qy        800 DVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856
Db        886 RVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942
```

RESULT 34
ABG73734
ID ABG73734 standard; protein; 584 AA.
XX
AC ABG73734;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 1 TH-1 variant TaHo C.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo C;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP; variant;
KW mutant.
XX
OS Synthetic.

```
XX
FH   Key             Location/Qualifiers
FT   Region          1. .22
FT                   /label= Ankyrin_repeat_1
FT   Region          24. .55
FT                   /label= Ankyrin_repeat_2
FT   Region          57. .87
FT                   /label= Ankyrin_repeat_3
FT   Region          89. .113
FT                   /label= Ankyrin_repeat_4
FT   Region          115. .139
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 127
FT                   /note= "TH1 start"
FT   Region          147. .176
FT                   /label= Ankyrin_repeat_6
FT   Region          178. .208
FT                   /label= Ankyrin_repeat_7
FT   Region          210. .241
FT                   /label= Ankyrin_repeat_8
FT   Region          243. .281
FT                   /label= Ankyrin_repeat_9
FT   Region          283. .332
FT                   /label= Ankyrin_repeat_10
FT   Region          334. .365
FT                   /label= Ankyrin_repeat_11
FT   Region          367. .398
FT                   /label= Ankyrin_repeat_12
FT   Region          400. .424
FT                   /label= Ankyrin_repeat_13
FT   Region          426. .458
FT                   /label= Ankyrin_repeat_14
FT   Region          460. .491
FT                   /label= Ankyrin_repeat_15
FT   Region          493. .524
FT                   /label= Ankyrin_repeat_16
FT   Region          526. .557
FT                   /label= Ankyrin_repeat_17
FT   Region          559. .584
FT                   /label= Ankyrin_repeat_18
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-C; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform TH-1 variant TaHo C which
CC   contains a C-terminus deletion and a 28 amino acid insertion due to a
CC   frame shift described in the disclosure of the invention
XX
SQ   Sequence 584 AA;

  Query Match            45.1%;  Score 2913;  DB 6;  Length 584;
  Best Local Similarity  99.8%;  Pred. No. 1.1e-250;
  Matches  554;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        681 ICKLLLQHGADPTKK 695
              |||||||||||||||
Db        541 ICKLLLQHGADPTKK 555


RESULT 35
AAB66287
ID   AAB66287 standard; protein; 522 AA.
XX
AC   AAB66287;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone protein sequence SEQ ID NO: 99.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
```

XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63925.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 156-157; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 522 AA;

Query Match 42.1%; Score 2722; DB 4; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.1e-233;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          72 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 131
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 60

Qy         132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 120

Qy         192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 251
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 180

Qy         252 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 311
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 240

Qy         312 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 371
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 300

Qy         372 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 360

Qy         432 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 491
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 420

Qy         492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 551
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 480

Qy         552 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 593
               |||||||||||||||||||||||||||||||||||||||||||
Db         481 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 522
```

RESULT 36
ABB11320
ID ABB11320 standard; peptide; 415 AA.
XX
AC ABB11320;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human Grb7 effector protein homologue, SEQ ID NO:1690.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA08564.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 20; Page 165-166; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides

```
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a novel human
CC   polypeptide of the invention
XX
SQ   Sequence 415 AA;

  Query Match             33.5%;  Score 2165;  DB 4;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-184;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        348 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 60

Qy        408 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 120

Qy        468 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 180

Qy        528 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 240

Qy        588 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 300

Qy        648 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 360

Qy        708 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 413


RESULT 37
AAY44403
ID   AAY44403 standard; protein; 673 AA.
XX
AC   AAY44403;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human truncated tankyrase-1.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   telomere length regulation; recombinant tankyrase; aging; skin atrophy;
KW   macular degeneration; atherosclerosis; ataxia telangiectasia;
KW   tankyrase modulator; rational drug design.
XX
OS   Homo sapiens.
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   N-PSDB; AAZ29632.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Disclosure; Page 121-123; 133pp; English.
XX
CC   The present sequence is truncated human tankyrase, a TRF1 (telomeric
CC   repeat binding factor-1) binding protein. This comprises 1-640 residues
CC   of human tankyrase. Tankyrase is involved in regulation of telomere
CC   length through TRF1 and may directly modulate the effect of TRF1.
CC   Recombinant tankyrase, or its fragments are used to identify specific
CC   modulators which are potential drugs for countering telomere shortening
CC   associated with aging (e.g. atrophy of skin, macular degeneration or
CC   atherosclerosis) or diseases like ataxia telangiectasia. They are also
CC   used to raise specific antibodies (used as immunoassay reagents and as
CC   modulators of tankyrase activity) and in rational drug design
XX
SQ   Sequence 673 AA;

  Query Match             33.0%;  Score 2132.5;  DB 3;  Length 673;
  Best Local Similarity   73.3%;  Pred. No. 8.6e-181;
  Matches  417;  Conservative   54;  Mismatches   91;  Indels    7;  Gaps    2;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :|| : : |:|     :   : ||      : |      | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||    ||  | ||| |||||||| ||||||    ||::| ||||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||:|||  ||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||:|||||||||||||||||||||||||||||||||||:|||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:|||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||||| ||:|||||||| ||||| |||:|
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYGS 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQS 590
              |    |   ||: |: | |::: :  |
Db        646 DPSIISLQGFTAAQMGN-EAVQQILSGHS 673


RESULT 38
AAU79538
ID   AAU79538 standard; protein; 673 AA.
XX
AC   AAU79538;
XX
DT   12-AUG-2002  (first entry)
XX
DE   Truncated human tankyrase #1.
```

```
XX
KW   Human; tankyrase; TRF1 interacting ankyrin; cytostatic; ADP ribosylation;
KW   telomere; nuclear localisation signal; NLS;
KW   telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW   drug development; aging; ataxia telangiectasia; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO200229017-A2.
XX
PD   11-APR-2002.
XX
PF   25-SEP-2001; 2001WO-US029931.
XX
PR   03-OCT-2000; 2000US-00678411.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2002-463230/49.
DR   N-PSDB; ABK48634.
XX
PT   Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT   to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT   ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT   -ribosylation activity.
XX
PS   Disclosure; Page 138-140; 151pp; English.
XX
CC   The invention discloses a nuclear localisation signal (NLS)-telomeric
CC   protein tankyrase that comprises a NLS and at least a portion of a
CC   tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC   factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC   Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC   TRF1, which is a negative regulator of telomere length maintenance, and
CC   in so doing, counteracts its effect and play a role in lifecycle of a
CC   cell. The protein is useful for identifying agents that modulate the
CC   binding of TRF1 to its DNA binding site, for identifying agents that
CC   modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC   identifying the agents that modulate telomere length and for identifying
CC   agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC   the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC   would hinder telomere lengthening and so is useful in the treatment of
CC   cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC   telomere lengthening and so could be useful for extending the lifetime of
CC   the cell, especially self-renewing cells such as bone marrow cells,
CC   peripheral blood, T and B cells. Tankyrases are also useful for
CC   developing drugs that can counteract the telomere shortening associated
CC   with aging and certain diseases such as ataxia telangiectasia. The
CC   protein sequence presented is the truncated human tankyrase #1. Note:
CC   Residues 1-640 are the same as the wild-type human tankyrase sequence
CC   (AAU79537)
XX
SQ   Sequence 673 AA;

  Query Match             33.0%;  Score 2132.5;  DB 5;  Length 673;
  Best Local Similarity   73.3%;  Pred. No. 8.6e-181;
  Matches  417;  Conservative   54;  Mismatches   91;  Indels    7;  Gaps    2;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |      |  :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||        ||  | ||| |||||||| ||||||    ||::| ||||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||||:|||  ||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||||:|||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              ||||||||||||||:|||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: ||||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||||| ||:|||||||| ||||| |||:|
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYGS 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQS 590
                |    |   ||: |: | |::: :  |
Db        646 DPSIISLQGFTAAQMGN-EAVQQILSGHS 673


RESULT 39
AAB66301
ID   AAB66301 standard; protein; 1099 AA.
XX
AC   AAB66301;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 expression plasmid protein sequence SEQ ID NO: 178.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63990.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 7; Page 236-239; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
```

```
SQ   Sequence 1099 AA;

  Query Match             31.8%;  Score 2056.5;  DB 4;  Length 1099;
  Best Local Similarity   62.1%;  Pred. No. 1.3e-173;
  Matches  459;  Conservative   38;  Mismatches  115;  Indels  127;  Gaps   17;

Qy        544 LQMGNENVQQLLQEGISLG---NSEADRQLLEAAKAGDVETV-KKLCTVQSVNCRDIEGR 599
Db        437 LSRNKDEVKAMIEK---LGGKLTGTANKASLCISTKKEVEKMNKKMEEVKEANIR----- 488

Qy        600 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 659
Db        489 --------------VVSEDFLQ---DVSASTK----------SLQELFLAHILSPWGAEV 521

Qy        660 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 719
Db        522 ------KAEPVEVVAPRGK---------SGAALSKKSK-GQVKEEGINKSEKRMKLTLKG 565

Qy        720 DAAL------------------LDAAKKGCLARVK--------KLSSPDNVN---CRDTQ 750
Db        566 GAAVDPDSGLEHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSW 625

Qy        751 GRHSTPLHLAAGYNNLE---VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 807
Db        626 GRVGTVI----GSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSKNFTKYPK--------- 672

Qy        808 YNACVNATDKWAFTPLH------EAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
Db        673 -----------KFYPLEIDYGQDEEAVKKLT----------VNPGTKRVEGQTPLDLVSA 711

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
Db        712 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 771

Qy        922 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 981
Db        772 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 831

Qy        982 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1041
Db        832 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 891

Qy       1042 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 1101
Db        892 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 951

Qy       1102 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1161
Db        952 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1011

Qy       1162 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1221
Db       1012 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1071

Qy       1222 YPEYLITYQIMRPEGMVDG 1240
Db       1072 YPEYLITYQIMRPEGMVDG 1090


RESULT 40
ABU70878
ID   ABU70878 standard; protein; 373 AA.
XX
AC   ABU70878;
XX
DT   10-JUN-2003  (first entry)
XX
DE   Human adipocyte Selected Interacting domain, SID, #509.
XX
KW   Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KW   antidiabetic; protein-protein interaction; diabetes;
KW   yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS   Homo sapiens.
XX
PN   WO200286122-A2.
XX
PD   31-OCT-2002.
XX
PF   14-MAR-2002; 2002WO-EP003768.
XX
PR   14-MAR-2001; 2001US-0275734P.
XX
PA   (HYBR-) HYBRIGENICS.
XX
PI   Legrain P,  Daviet L;
XX
DR   WPI; 2003-103412/09.
DR   N-PSDB; ACA57422.
XX
PT   New complex between two interacting proteins in adipocyte cells, useful
PT   for identifying selected interacting domains that modulate protein
PT   interactions, or for preventing or treating metabolic disorders such as
PT   obesity or diabetes.
XX
PS   Claim 6; Page 276; 382pp; English.
XX
CC   The invention relates to a complex between two interacting proteins in
CC   adipocyte cells, given in the specification. The proteins are identified
CC   by selecting a bait protein from a known adipocyte marker and then
CC   performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC   members of an adipocyte cDNA library. The proteins are designated SID
CC   (RTM) (selected interacting domains) proteins. Also included are a
CC   polynucleotide encoding a polypeptide in the adipocyte cells, a
CC   recombinant host cell expressing at least one of the interacting
CC   polypeptides of the complex, selecting a modulating compound in adipocyte
CC   cells, a SID (RTM)) polypeptide comprising any of the 738 amino acid
CC   sequences given in the specification (including its fragment or variant),
CC   a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC   given in the specification (including its fragment or variant), a vector
CC   comprising the SID (RTM) polynucleotide, a recombinant host cell
CC   comprising the vector, a protein chip comprising the polypeptides and a
CC   record comprising all or part of the data, listed in the specification.
CC   The complex, polypeptides, polynucleotides and compounds are useful for
CC   preventing or treating metabolic disorders such as obesity or diabetes.
CC   The polynucleotides are useful as probes or primers. The complex is
CC   particularly useful for identifying selected interacting domains (SID
CC   (RTM)) for screening drugs that modulate the protein interaction, thus
CC   exhibiting the therapeutic effect. The present sequence represents a SID
CC   (prey) protein of the invention
XX
SQ   Sequence 373 AA;

  Query Match             30.4%;  Score 1967;  DB 6;  Length 373;
  Best Local Similarity   100.0%;  Pred. No. 2e-166;
  Matches  373;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        153 QNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKID 212
Db          1 QNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKID 60

Qy        213 VCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 272
Db         61 VCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 120

Qy        273 PLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYE 332
Db        121 PLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYE 180

Qy        333 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 392
Db        181 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 240

Qy        393 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 452
Db        241 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 300
```

```
Qy        453 RKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 RKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 360

Qy        513 RAAYCGHLQTCRL 525
              |||||||||||||
Db        361 RAAYCGHLQTCRL 373
```

RESULT 41
AAB48571
ID AAB48571 standard; protein; 385 AA.
XX
AC AAB48571;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human T cell leukaemia HOM-TALL1-5 polypeptide.
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; HOM-TALL1-5.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR N-PSDB; AAC89701, AAC89702.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 1; Page 75-76; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 385 AA;

```
  Query Match             30.0%;  Score 1942;  DB 4;  Length 385;
  Best Local Similarity   99.5%;  Pred. No. 3.5e-164;
  Matches  369;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        177 HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALD 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALD 60

Qy        237 LADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYN 296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LADPSAKAVLTGEYKKDELLESARXGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYN 120

Qy        297 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH 180

Qy        357 EAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREA 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREA 240

Qy        417 DVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 300

Qy        477 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNII 536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLHTCRLLLSYGCDPNII 360

Qy        537 SLQGFTALQMG 547
              |||||||||||
Db        361 SLQGFTALQMG 371
```

RESULT 42
AAB66300
ID AAB66300 standard; protein; 362 AA.
XX
AC AAB66300;
XX
DT 05-APR-2001 (first entry)
XX
DE Human TRF1 TANK2 binding domain protein sequence SEQ ID NO: 160.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63973.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 6; Page 228-229; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 362 AA;

```
  Query Match             29.3%;  Score 1895;  DB 4;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-160;
```

```
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        352 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 60

Qy        412 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 120

Qy        472 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 180

Qy        532 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 240

Qy        592 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 300

Qy        652 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 360

Qy        712 DI 713
              ||
Db        361 DI 362


RESULT 43
ABG73733
ID   ABG73733 standard; protein; 338 AA.
XX
AC   ABG73733;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TH-2.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TH-2; TaHo;
KW   cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Region          141. .162
FT                   /label= Ankyrin_repeat_1
FT   Region          164. .195
FT                   /label= Ankyrin_repeat_2
FT   Region          197. .227
FT                   /label= Ankyrin_repeat_3
FT   Region          229. .253
FT                   /label= Ankyrin_repeat_4
FT   Region          255. .279
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 267
FT                   /note= "TH1 start"
FT   Region          287. .316
FT                   /label= Ankyrin_repeat_6
FT   Region          318. .338
FT                   /label= Ankyrin_repeat_7
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-B; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents tankyrase H isoform 1, TH-2, described in the
CC   disclosure of the invention
XX
SQ   Sequence 338 AA;

  Query Match             26.9%;  Score 1736;  DB 6;  Length 338;
  Best Local Similarity   100.0%;  Pred. No. 7.4e-146;
  Matches  338;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240

Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLV 338
              ||||||||||||||||||||||||||||||||||||||
Db        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLV 338


RESULT 44
AAB48572
ID   AAB48572 standard; protein; 331 AA.
XX
AC   AAB48572;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Human breast cancer MO-BC-203 polypeptide.
XX
KW   Human; cytostatic; gene therapy; vaccine; breast cancer;
KW   T cell leukaemia; cancer associated antigen; MO-BC-203.
XX
OS   Homo sapiens.
XX
PN   WO200072021-A2.
XX
```

PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR N-PSDB; AAC89703, AAC89704.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 1; Page 78-79; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 331 AA;

Query Match 26.4%; Score 1706; DB 4; Length 331;
Best Local Similarity 97.9%; Pred. No. 3.4e-143;
Matches 324; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy        337 LVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQL 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQL 60

Qy        397 KERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQI 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQI 120

Qy        457 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 516
              |||||:||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELLLKKGSNINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 180

Qy        517 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 240

Qy        577 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
              |||:||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        241 GDVKTVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKEKGGLVP 300

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWKF 667
              |||||||||||||||||||| || :||||
Db        301 LHNACSYGHYEVAELLVKHGAVFNVTELWKF 331
```

RESULT 45
ABU70889
ID ABU70889 standard; protein; 308 AA.
XX
AC ABU70889;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human adipocyte Selected Interacting domain, SID, #520.
XX
KW Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KW antidiabetic; protein-protein interaction; diabetes;
KW yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS Homo sapiens.
XX
PN WO200286122-A2.
XX
PD 31-OCT-2002.
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Daviet L;
XX
DR WPI; 2003-103412/09.
DR N-PSDB; ACA57433.
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
PS Claim 6; Page 279; 382pp; English.
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM)) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence represents a SID
CC (prey) protein of the invention
XX
SQ Sequence 308 AA;

Query Match 22.7%; Score 1467; DB 6; Length 308;
Best Local Similarity 88.3%; Pred. No. 7e-122;
Matches 272; Conservative 23; Mismatches 13; Indels 0; Gaps 0;

```
Qy        158 VQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVL 217
              |||||||||||||||||||||||:|||  ||||||||||||||||||||||||||||||
Db          1 VHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTPLHEAAIKGKIDVCIVL 60

Qy        218 LQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVN 277
              |||||:| ||||||::||||||||||||||||||||:||||||||:||||||||||
Db         61 LQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARSGNEEKLMALLTPLNVN 120

Qy        278 CHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELL 337
              ||||||||||||||||||:|||||||||||||||| ||||||||||||||||||
Db        121 CHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVPLHNACSYGHYEVTELL 180
```

```
Qy        338 VKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLK 397
              :|||||||||||||||||||||||||||||||||||:||||||:||| |||:|:||||:|:
Db        181 LKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVNCHGKSAVDMAPTPELR 240

Qy        398 ERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQIC 457
              ||| |||||||||||||||: ::|| |:||::||| ||:|||||||| || :|||||:
Db        241 ERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETALHCAVASLHPKRKQVT 300

Qy        458 ELLLRKGA 465
              ||||||||
Db        301 ELLLRKGA 308


Search completed: March 25, 2004, 16:41:38
Job time : 77.8205 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 16:15:57 ; Search time 1364.9 Seconds
(without alignments)
11877.169 Million cell updates/sec

Title: US-09-843-159B-2
Perfect score: 3816
Sequence: 1 cgcgctgctccgcccgccgc.....cctttgctgaaaaaaaaaaa 3816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3816	100.0	3816	4	AAA91488	Aaa91488 Tankyrase
2	3816	100.0	3816	7	ABQ77067	Abq77067 Tankyrase
3	3793	99.4	6019	4	AAC85294	Aac85294 Human SPA
4	3791.4	99.4	4275	3	AAC66825	Aac66825 Human tan
5	3791.4	99.4	4406	5	AAF63930	Aaf63930 Human tan
6	3791.4	99.4	4992	5	AAF63953	Aaf63953 Human tan
7	3791.4	99.4	5002	5	AAF63952	Aaf63952 Human tan
8	3788.4	99.3	6018	7	AAD54031	Aad54031 Human col
9	3770.8	98.8	5075	6	ABZ11674	Abz11674 Human pol
10	3595.8	94.2	4127	5	AAF63926	Aaf63926 Human tan
11	3594.8	94.2	4512	4	AAD02578	Aad02578 Human tan
12	3504.8	91.8	3508	5	AAF63837	Aaf63837 Human tan
13	3393.4	88.9	3797	4	AAA91487	Aaa91487 Tankyrase
14	3393.4	88.9	3797	7	ABQ77066	Abq77066 Tankyrase
15	3341	87.6	4296	3	AAC66824	Aac66824 Human tan
16	3297	86.4	3400	2	AAX25366	Aax25366 Human Grb
17	3039	79.6	4493	3	AAC66823	Aac66823 Human tan
18	2774.6	72.7	5005	4	AAC89706	Aac89706 Human adu
19	2439.8	63.9	2971	5	AAF63919	Aaf63919 Human tan
20	2439.8	63.9	3353	5	AAF63917	Aaf63917 Human tan
21	2357.8	61.8	3799	5	AAF63918	Aaf63918 Human tan
22	2122	55.6	3254	4	AAC85295	Aac85295 Mouse SPA
23	1885	49.4	3984	5	AAF63838	Aaf63838 Human tan
24	1885	49.4	4134	3	AAZ29627	Aaz29627 Human tan
25	1885	49.4	4134	3	AAC66826	Aac66826 Human tan
26	1885	49.4	4134	6	ABK48629	Abk48629 Human tan
27	1568	41.1	1568	5	AAF63925	Aaf63925 Human tan
28	1533.8	40.2	4435	9	ADE53995	Ade53995 Human pro
29	1521.2	39.9	4491	3	AAZ29632	Aaz29632 Human tru
30	1521.2	39.9	4491	6	ABK48634	Abk48634 Truncated
31	1369.4	35.9	4657	3	AAZ29633	Aaz29633 Human tru
32	1369.4	35.9	4657	6	ABK48635	Abk48635 Truncated
33	1240.4	32.5	1466	4	ABA08564	Aba08564 Human Grb
34	1192.6	31.3	2409	6	AAK99406	Aak99406 DNA of AP
35	1172.2	30.7	1242	4	AAC89701	Aac89701 Human adu
36	1168.4	30.6	3308	5	AAF63990	Aaf63990 Human tan
37	1121	29.4	1121	7	ACA57422	Aca57422 Human adi
38	1084.8	28.4	1103	5	AAF63973	Aaf63973 Human TRF
39	1076.8	28.2	3006	4	AAH16624	Aah16624 Human cDN
40	1068.4	28.0	5484	4	ABL08155	Abl08155 Drosophil
41	1066.8	28.0	5482	5	AAF63955	Aaf63955 Drosophil
42	990	25.9	1007	4	AAC89703	Aac89703 Human adu
43	794.4	20.8	797	9	ADE53994	Ade53994 Human pro
44	786	20.6	1344	3	AAC77578	Aac77578 Human ORF
45	784.2	20.6	838	4	AAH08231	Aah08231 Human cDN

ALIGNMENTS

RESULT 1
AAA91488
ID AAA91488 standard; DNA; 3816 BP.
XX
AC AAA91488;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 2 coding sequence.
XX
KW Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2. .3726
FT /*tag= a
FT /partial
FT /product= "TaHo2"
FT /note= "Tankyrase homologue isotype 2; No start codon
FT given"
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR P-PSDB; AAY97749.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 13; Fig 2; 63pp; English.
XX
CC This sequence encodes the Tankyrase homologue isotype 2 (TaHo-2) protein

CC of the invention. The invention also relates to the TaHo-1 protein. The
CC TaHo proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as
CC vaccine, and for construction of hybridisation probes for mapping the
CC gene which encodes TaHo and for the genetic analysis of individuals with
CC genetic disorders. The TaHo proteins, and their coding sequences are
CC useful in screening assays
XX
SQ Sequence 3816 BP; 1097 A; 806 C; 949 G; 964 T; 0 U; 0 Other;

Query Match 100.0%; Score 3816; DB 4; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240

Qy        241 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260

Qy       1261 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560

Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680

Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740

Qy       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800

Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860

Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
```

```
Db       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Qy       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Qy       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Qy       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Qy       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
Db       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
```

RESULT 2
ABQ77067
ID ABQ77067 standard; DNA; 3816 BP.
XX
AC ABQ77067;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 2 TaHo-2 DNA from clone K23.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-2;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP; ds.
XX

OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 3. .3725
FT /*tag= a
FT /product= "TaHo-2"
FT /partial
FT /note= "no start codon"
XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
DR P-PSDB; ABG73731.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Claim 13; Fig 2A-B; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence encodes the tankyrase H isoform 2, TaHo-2, isolated from clone
CC K23 which is described in the disclosure of the invention
XX
SQ Sequence 3816 BP; 1097 A; 806 C; 949 G; 964 T; 0 U; 0 Other;

Query Match 100.0%; Score 3816; DB 7; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200

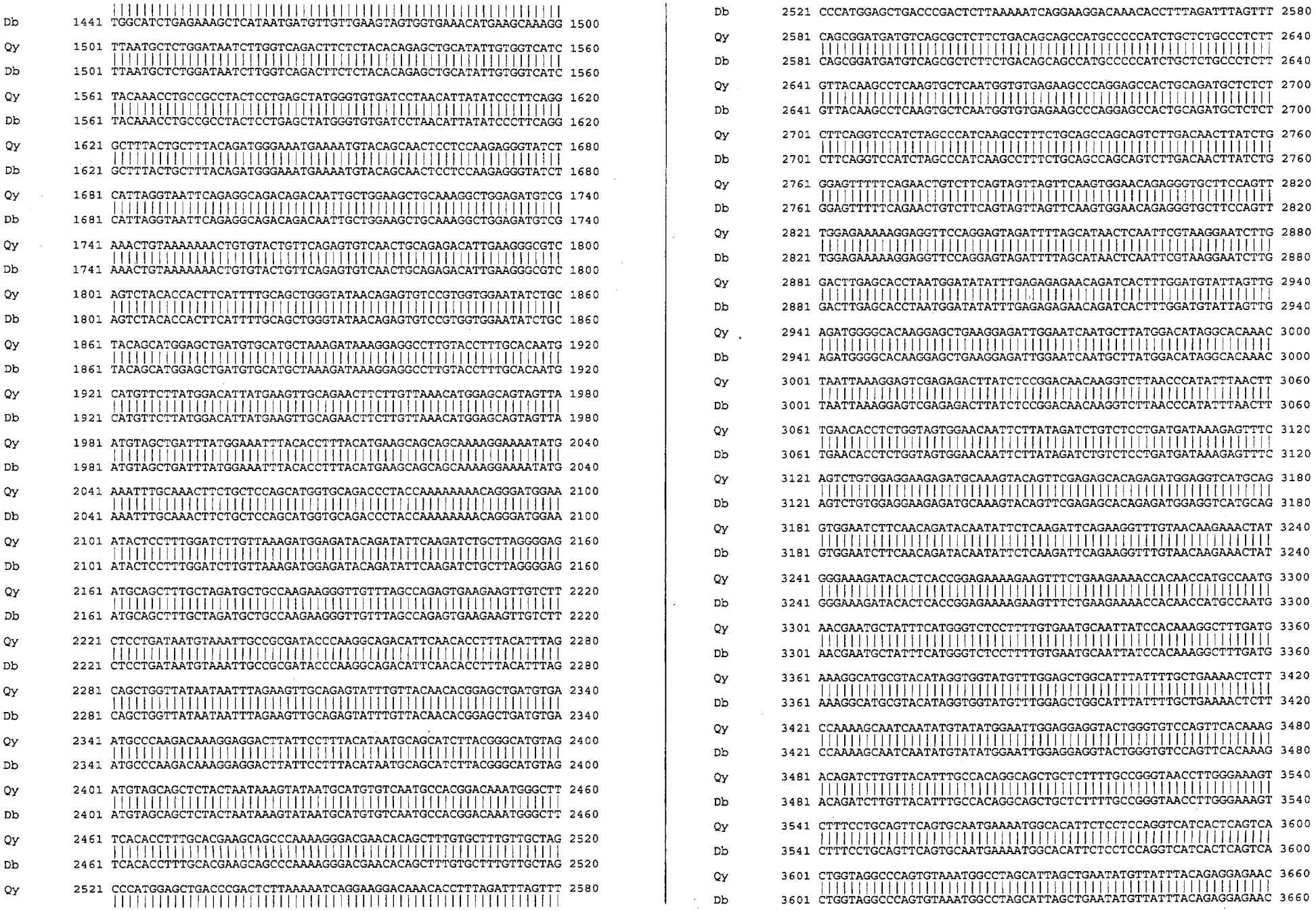
Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
```


Db 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Qy 2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db 2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Qy 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Qy 2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db 2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Qy 2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db 2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Qy 2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db 2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Qy 2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db 2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Qy 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Qy 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Qy 2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db 2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Qy 3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db 3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Qy 3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db 3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Qy 3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db 3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Qy 3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db 3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Qy 3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db 3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Qy 3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db 3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Qy 3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db 3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Qy 3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db 3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Qy 3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db 3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Qy 3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db 3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Qy 3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db 3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660

```
Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
             ||||||||||||||||||||||||||||||||||||
Db       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
```

```
RESULT 3
AAC85294
ID   AAC85294 standard; cDNA; 6019 BP.
XX
AC   AAC85294;
XX
DT   29-MAR-2001  (first entry)
XX
DE   Human SPANK cDNA.
XX
KW   SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW   poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW   ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW   adipocyte; insulin signalling pathway; hyperlipidaemia;
KW   glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW   cardiac insufficiency; coronary insufficiency; stroke;
KW   high blood pressure; non-insulin dependent diabetes; hypertension;
KW   hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             252. .3735
FT                   /*tag= a
FT                   /product= "Human SPANK"
FT                   /transl_except= (pos:768. .878,aa:173. .209)
FT                   /transl_except= (pos:1230. .1337,aa:327. .362)
FT                   /transl_except= (pos:1374. .1445,aa:375. .398)
FT                   /transl_except= (pos:1695. .1823,aa:482. .524)
FT                   /transl_except= (pos:2172. .2282,aa:641. .677)
XX
PN   WO200077225-A1.
XX
PD   21-DEC-2000.
XX
PF   09-JUN-2000; 2000WO-US015926.
XX
PR   11-JUN-1999;   99US-0138957P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Chi N,  Lodish HF;
XX
DR   WPI; 2001-091404/10.
DR   P-PSDB; AAB47022.
XX
PT   New insulin signaling protein SPANK, useful for reducing body mass,
PT   glucose intolerance or insulin resistance and for preventing or treating
PT   obesity-related and muscle-related diseases.
XX
PS   Claim 3; Fig 4; 65pp; English.
XX
CC   This sequence represents the human SPANK cDNA. The SPANK protein
CC   comprises 3 domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP
CC   (poly adenosine diphosphate-ribose polymerase) catalytic domain; and (c)
CC   an ANK domain composed of ankyrin repeats. SPANK is a cytosolic protein
CC   which can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC   aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC   perinuclear region of adipocytes. It is an effector in the insulin
CC   signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC   mass, reducing glucose intolerance or insulin resistance, for preventing
CC   or treating obesity-related diseases or disorders, such as obesity,
CC   cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC   atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC   dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC   also useful for preventing or treating muscle-related diseases or
CC   disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC   Antibodies immunospecific for SPANK are useful for detecting the presence
CC   of SPANK polypeptide in a biological sample
XX
SQ   Sequence 6019 BP; 1732 A; 1173 C; 1341 G; 1773 T; 0 U; 0 Other;
```

```
  Query Match           99.4%;  Score 3793;  DB 4;  Length 6019;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 3807;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 88

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 147

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 207

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 267

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 327

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 387

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 447

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 507

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 567

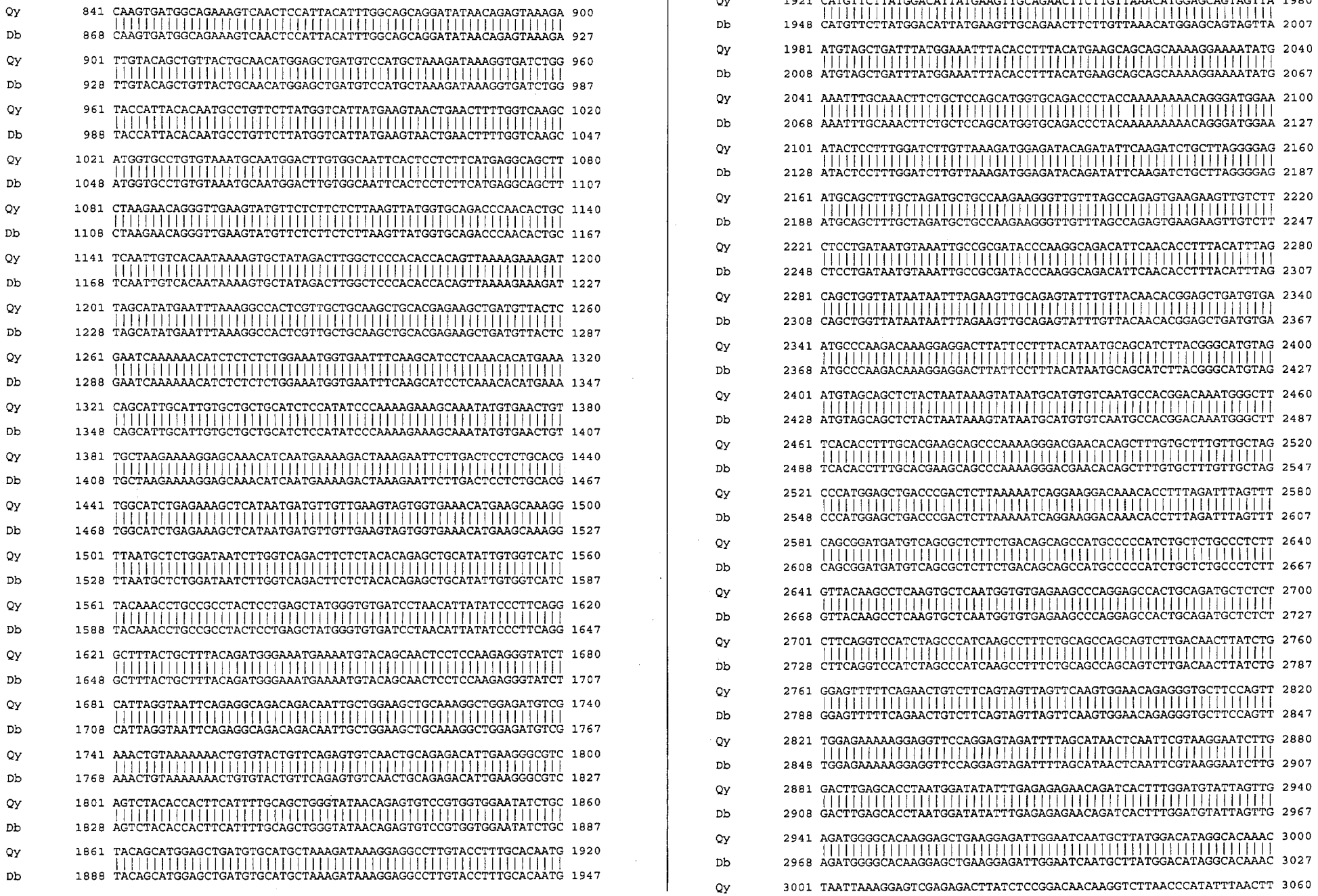
Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 627

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 687

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 747

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 807

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 867
```


Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 868 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 927
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 928 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 987
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 988 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1047
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1048 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1107
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1108 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1167
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1168 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1227
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1228 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1287
Qy 1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1288 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1347
Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1348 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1407
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 1408 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1467
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 1468 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1527
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 1528 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1587
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 1588 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1647
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 1648 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1707
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 1708 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1767
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1768 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1827
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 1828 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1887
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 1888 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1947
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 1948 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2007
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2008 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2067
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db 2068 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2127
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2128 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2187
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2188 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2247
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2248 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2307
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db 2308 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2367
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 2368 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2427
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db 2428 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2487
Qy 2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db 2488 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2547
Qy 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 2548 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2607
Qy 2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db 2608 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2667
Qy 2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db 2668 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2727
Qy 2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db 2728 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2787
Qy 2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db 2788 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2847
Qy 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 2848 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2907
Qy 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db 2908 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2967
Qy 2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db 2968 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3027
Qy 3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060

```
Db       3028 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3087

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120

Db       3088 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3147

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180

Db       3148 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3207

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240

Db       3208 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3267

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300

Db       3268 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3327

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360

Db       3328 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3387

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420

Db       3388 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3447

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480

Db       3448 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3507

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540

Db       3508 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3567

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600

Db       3568 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3627

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660

Db       3628 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3687

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720

Db       3688 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3747

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780

Db       3748 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3807

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813

Db       3808 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3840
```

RESULT 4
AAC66825
ID AAC66825 standard; cDNA; 4275 BP.
XX
AC AAC66825;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase II coding sequence SEQ ID NO: 5.
XX
KW Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 284. .3784
FT /*tag= a
FT /product= "tankyrase II"
FT /transl_except= (pos:655. .660,aa:TyrThr)
XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009558.
XX
PR 09-APR-1999; 99US-0128577P.
PR 13-APR-1999; 99US-0129123P.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
DR P-PSDB; AAB27211.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 1; Fig 4; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase II coding
CC sequence. Its protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell
XX
SQ Sequence 4275 BP; 1247 A; 889 C; 1025 G; 1114 T; 0 U; 0 Other;

Query Match 99.4%; Score 3791.4; DB 3; Length 4275;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Db         61 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAGGGGCGCGCG 120

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Db        121 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 179

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Db        180 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 239

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240

Db        240 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 299

Qy        241 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300

Db        300 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 359

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360

Db        360 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 419

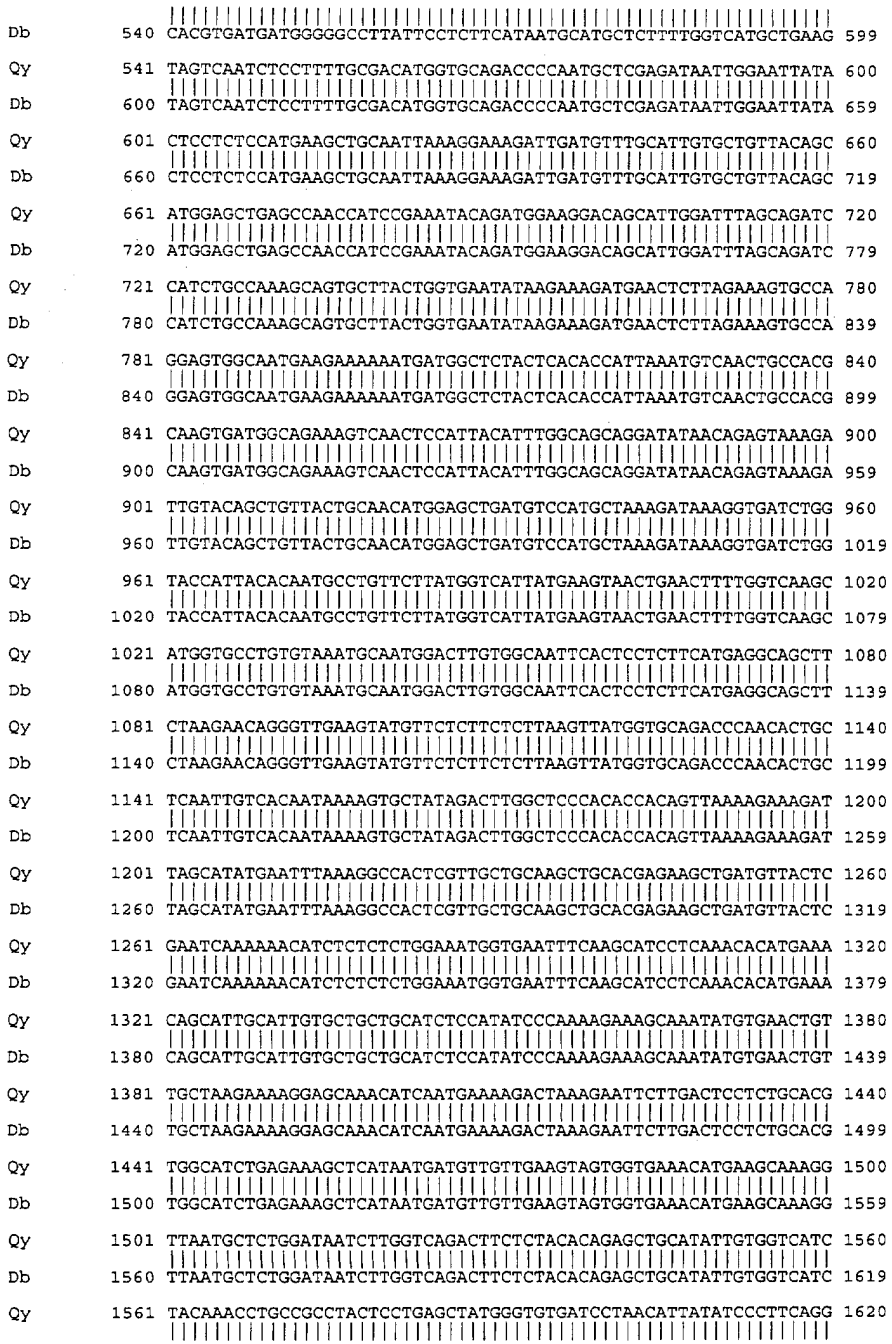
Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420

Db        420 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 479

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480

Db        480 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 539

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
```


Db 540 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 599
Qy 541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db 600 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 659
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db 660 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 719
Qy 661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db 720 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 779
Qy 721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db 780 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 839
Qy 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db 840 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 899
Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 900 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 959
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 960 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1019
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1020 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1079
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1080 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1139
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1140 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1199
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1200 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1259
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1260 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1319
Qy 1261 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1320 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1379
Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1380 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1439
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 1440 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1499
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 1500 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1559
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 1560 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1619
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 1620 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1679
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 1680 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1739
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 1740 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1799
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1800 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1859
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 1860 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1919
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 1920 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1979
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 1980 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2039
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2040 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2099
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db 2100 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2159
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2160 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2219
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2220 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2279
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2280 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2339
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db 2340 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2399
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 2400 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2459
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db 2460 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2519
Qy 2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db 2520 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2579
Qy 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 2580 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2639
Qy 2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db 2640 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2699
Qy 2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db 2700 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2759

```
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2760 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2819

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2820 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2879

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2880 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2939

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2940 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2999

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3000 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3059

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3060 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3119

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3179

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3239

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3299

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3359

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3419

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3479

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3539

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3599

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3659

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3660 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3719

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3720 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3779

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3780 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3839
```

```
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             |||||||||||||||||||||||||||||||||
Db       3840 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 5
AAF63930
ID AAF63930 standard; DNA; 4406 BP.
XX
AC AAF63930;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 106.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66290.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 2; Page 168-173; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 4406 BP; 1268 A; 910 C; 1054 G; 1174 T; 0 U; 0 Other;

Query Match 99.4%; Score 3791.4; DB 5; Length 4406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 126

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 185

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 245

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 305

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
```

```
Db        306 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 365
Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db        366 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 425
Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db        426 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 485
Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db        486 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 545
Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db        546 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 605
Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db        606 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 665
Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db        666 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 725
Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db        726 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 785
Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db        786 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 845
Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db        846 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 905
Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db        906 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 965
Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db        966 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1025
Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db       1026 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1085
Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db       1086 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1145
Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db       1146 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1205
Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db       1206 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1265
Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db       1266 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1325
Qy       1261 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db       1326 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1385
Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
```

```
Db       1386 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1445
Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db       1446 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1505
Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db       1506 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1565
Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db       1566 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1625
Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db       1626 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1685
Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db       1686 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1745
Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db       1746 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1805
Qy       1741 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db       1806 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1865
Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db       1866 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1925
Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db       1926 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1985
Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db       1986 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2045
Qy       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db       2046 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2105
Qy       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db       2106 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2165
Qy       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db       2166 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2225
Qy       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db       2226 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2285
Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db       2286 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2345
Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db       2346 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2405
Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2406 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2465
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2466 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2525
```

```
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2526 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2585

Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2586 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2645

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2646 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2705

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2706 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2765

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2766 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2825

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2826 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2885

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2886 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2945

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2946 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3005

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3006 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3065

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3066 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3125

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3126 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3185

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3186 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3245

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3246 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3305

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3306 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3365

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3366 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3425

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3426 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3485

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3486 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3545

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3546 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3605

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3606 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3665

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3666 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3725

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3726 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3785

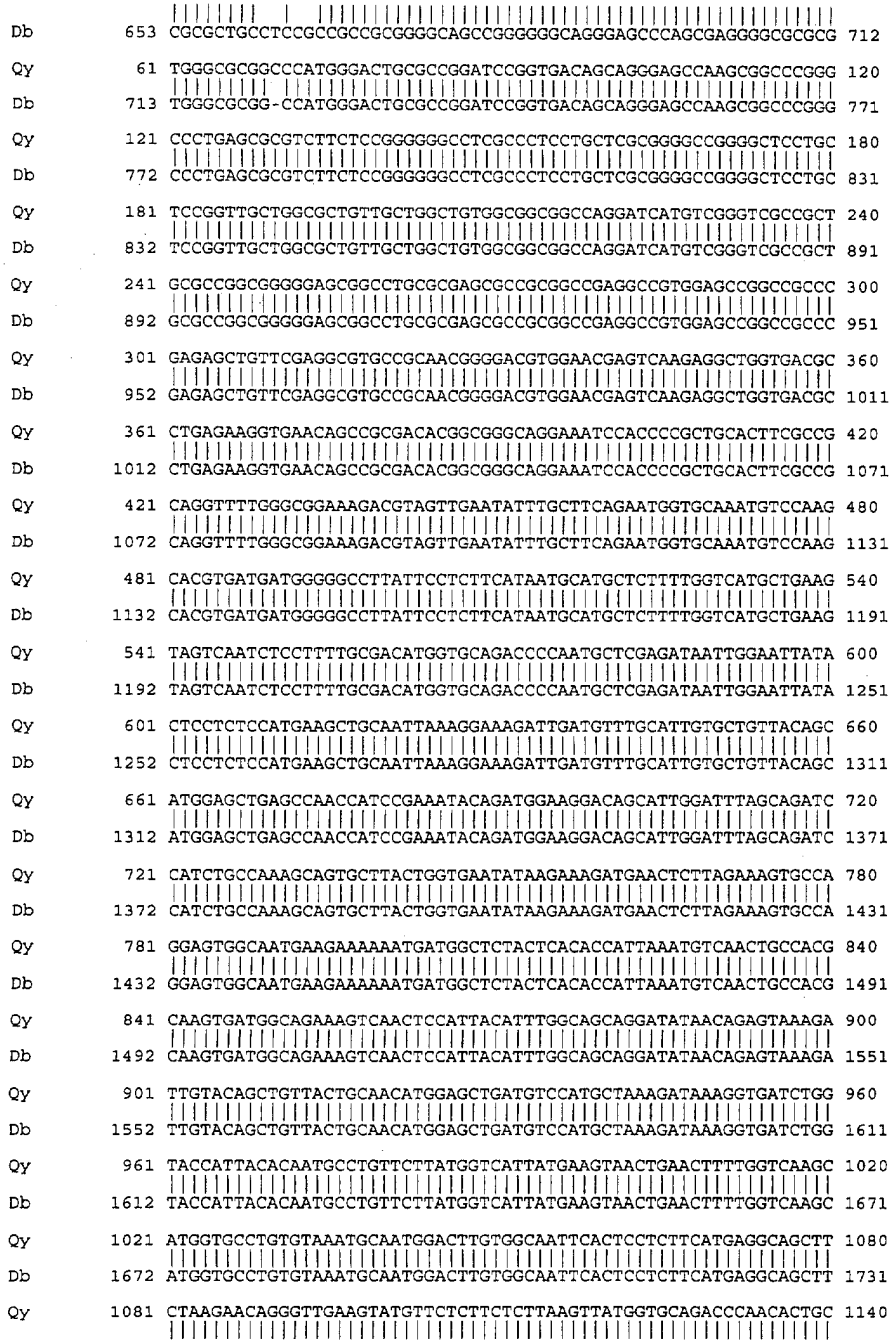
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3786 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3845

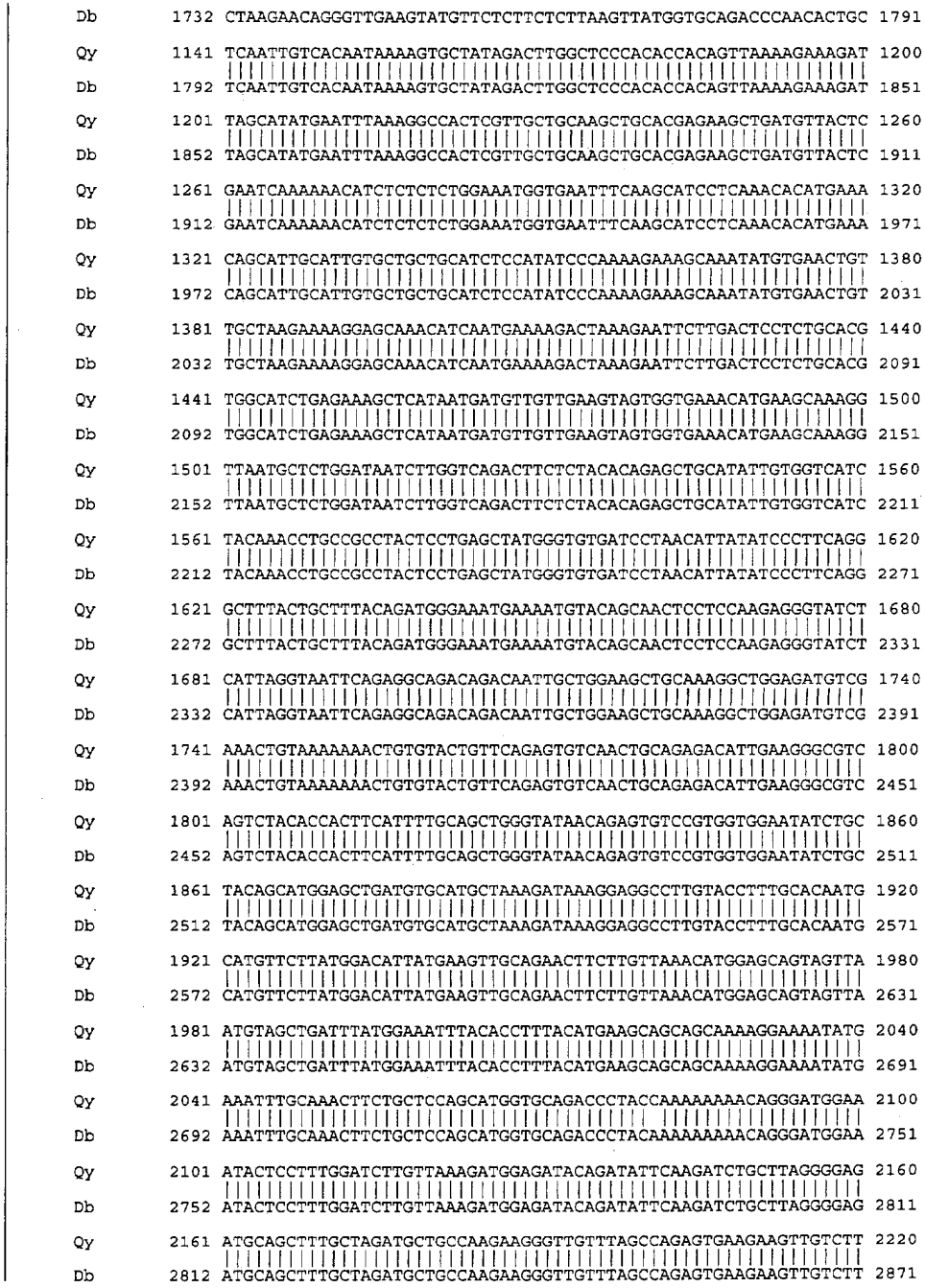
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||
Db       3846 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3878
```

RESULT 6
AAF63953
ID AAF63953 standard; DNA; 4992 BP.
XX
AC AAF63953;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-SHORT coding sequence SEQ ID NO: 134.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66295.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Claim 6; Page 195-200; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 4992 BP; 1454 A; 1064 C; 1191 G; 1283 T; 0 U; 0 Other;

Query Match 99.4%; Score 3791.4; DB 5; Length 4992;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
```


Db 653 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 712
Qy 61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
Db 713 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 771
Qy 121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
Db 772 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 831
Qy 181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
Db 832 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 891
Qy 241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db 892 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 951
Qy 301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db 952 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 1011
Qy 361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db 1012 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 1071
Qy 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db 1072 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 1131
Qy 481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db 1132 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 1191
Qy 541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db 1192 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 1251
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db 1252 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 1311
Qy 661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db 1312 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 1371
Qy 721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db 1372 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 1431
Qy 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db 1432 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 1491
Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 1492 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1551
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 1552 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1611
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1612 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1671
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1672 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1731
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140


Db 1732 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1791
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1792 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1851
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1852 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1911
Qy 1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1912 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1971
Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1972 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 2031
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 2032 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 2091
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 2092 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 2151
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 2152 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 2211
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 2212 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 2271
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 2272 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 2331
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 2332 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 2391
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 2392 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 2451
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 2452 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2511
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 2512 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2571
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 2572 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2631
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2632 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2691
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db 2692 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2751
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2752 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2811
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2812 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2871

```
Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2872 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2931

Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2932 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2991

Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2992 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3051

Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3052 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 3111

Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3112 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 3171

Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3172 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3231

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3232 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3291

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3292 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 3351

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3352 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 3411

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3412 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 3471

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3472 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3531

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3532 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3591

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3592 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3651

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3652 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3711

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3712 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3771

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3772 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3831

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3832 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3891

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3892 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3951

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3952 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 4011

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4012 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 4071

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4072 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 4131

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4132 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 4191

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4192 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 4251

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4252 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 4311

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4312 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 4371

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4372 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 4431

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||
Db       4432 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4464
```

```
RESULT 7
AAF63952
ID   AAF63952 standard; DNA; 5002 BP.
XX
AC   AAF63952;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 TANK2-LONG coding sequence SEQ ID NO: 132.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66294.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Claim 5; Page 185-190; 242pp; English.
XX
```

CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 5002 BP; 1456 A; 1065 C; 1196 G; 1285 T; 0 U; 0 Other;

Query Match 99.4%; Score 3791.4; DB 5; Length 5002;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 722

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 781

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 841

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 901

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 961

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 1021

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 1081

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 1141

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 1201

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 1261

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 1321

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 1381

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 1441

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 1501

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1502 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1561

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1621

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1681

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1741

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1801

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1861

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1921

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1981

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 2041

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 2101

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 2161

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 2221

Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 2281

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 2341

Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 2401

Qy       1741 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 2461

Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2521

Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2522 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2581

Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2582 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2641
```

```
Qy      1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2642 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2701

Qy      2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db      2702 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2761

Qy      2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2762 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2821

Qy      2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2822 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2881

Qy      2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2882 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2941

Qy      2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2942 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 3001

Qy      2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3002 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3061

Qy      2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3062 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 3121

Qy      2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3122 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 3181

Qy      2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3182 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3241

Qy      2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3242 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3301

Qy      2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3302 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 3361

Qy      2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3362 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 3421

Qy      2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3422 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 3481

Qy      2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3482 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3541

Qy      2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3542 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3601

Qy      2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3602 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3661

Qy      3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3662 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3721

Qy      3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3722 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3781

Qy      3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3782 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3841

Qy      3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3842 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3901

Qy      3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3902 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3961

Qy      3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3962 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 4021

Qy      3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4022 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 4081

Qy      3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4082 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 4141

Qy      3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4142 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 4201

Qy      3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4202 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 4261

Qy      3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4262 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 4321

Qy      3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4322 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 4381

Qy      3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4382 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 4441

Qy      3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||
Db      4442 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4474
```

RESULT 8
AAD54031
ID AAD54031 standard; DNA; 6018 BP.
XX
AC AAD54031;
XX
DT 17-JUN-2003 (first entry)
XX
DE Human colon cancer-associated polypeptide gene, TNKL (BC-203).
XX
KW Human; colon cancer-associated polypeptide; immune response; therapy;
KW colon cancer; gene; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	33. .3283
FT		/*tag= a
FT		/product= "Human colon cancer-associated polypeptide"
FT		/transl_except= (pos:84. .87, aa:Ser)

```
FT                   /transl_except= (pos:262. .265, aa:Asp)
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200290986-A1.
XX
PD   14-NOV-2002.
XX
PF   02-MAY-2002; 2002WO-US013994.
XX
PR   04-MAY-2001; 2001US-00849602.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chen Y,  Old LJ,  Scanlan MJ,  Stockert E;
XX
DR   WPI; 2003-112003/10.
DR   P-PSDB; AAE35352.
XX
PT   Diagnosing colon cancer in a subject comprises identifying colon cancer-
PT   associated polypeptides as antigens that elicit immune response in colon
PT   cancer.
XX
PS   Claim 1; Page 85-89; 122pp; English.
XX
CC   The invention relates to a method for diagnosing colon cancer in a
CC   subject which comprises identifying colon cancer-associated polypeptides
CC   as antigens that elicit immune response in colon cancer. The method is
CC   useful for diagnosing, determining onset, progression, or regression of
CC   colon cancer in a subject, or for selecting a course of treatment of a
CC   subject having or suspected of having colon cancer. The colon cancer-
CC   associated polypeptides are useful as markers for diagnosing colon
CC   cancer, and for following the course of treatment of colon cancer. The
CC   present sequence is human colon cancer-associated polypeptide gene
XX
SQ   Sequence 6018 BP; 1741 A; 1171 C; 1328 G; 1778 T; 0 U; 0 Other;

  Query Match             99.3%;  Score 3788.4;  DB 7;  Length 6018;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3803;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          4 GCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCGTGG 63
              |||||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCGTGG 60

Qy         64 GCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCC 123
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCC 119

Qy        124 TGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCC 179

Qy        184 GGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCG 239

Qy        244 CCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAG 299

Qy        304 AGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 AGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTG 359

Qy        364 AGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 AGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAG 419

Qy        424 GTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCAC 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 GTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCAC 479

Qy        484 GTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAG 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAG 539

Qy        544 TCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTC 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 TCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTC 599

Qy        604 CTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATG 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 CTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATG 659

Qy        664 GAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCAT 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 GAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCAT 719

Qy        724 CTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 CTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGA 779

Qy        784 GTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 GTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAA 839

Qy        844 GTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTG 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 GTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTG 899

Qy        904 TACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTAC 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 TACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTAC 959

Qy        964 CATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATG 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 CATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATG 1019

Qy       1024 GTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTA 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 GTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTA 1079

Qy       1084 AGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCA 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 AGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCA 1139

Qy       1144 ATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAG 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 ATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAG 1199

Qy       1204 CATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAA 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 CATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAA 1259

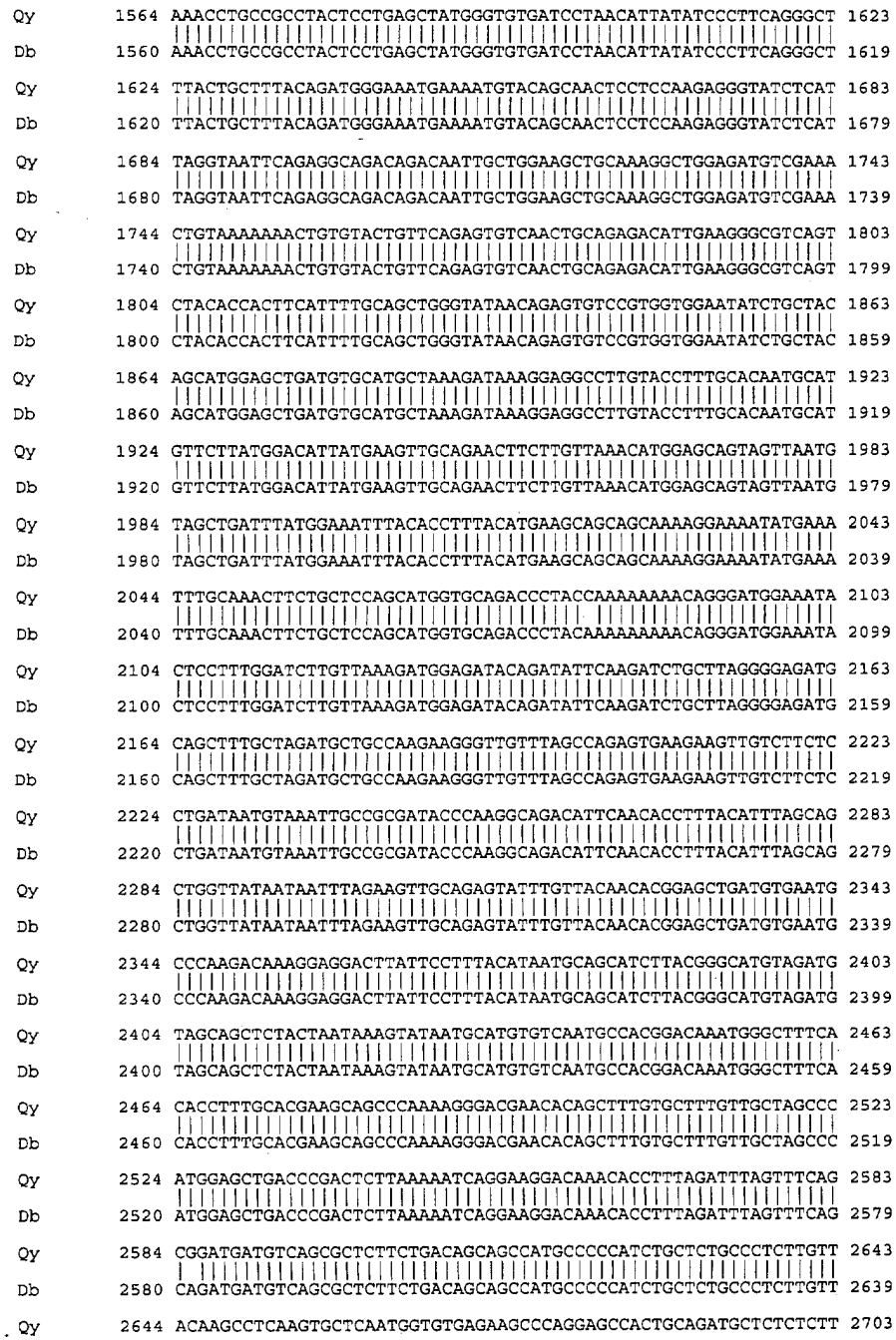
Qy       1264 TCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAG 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 TCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAG 1319

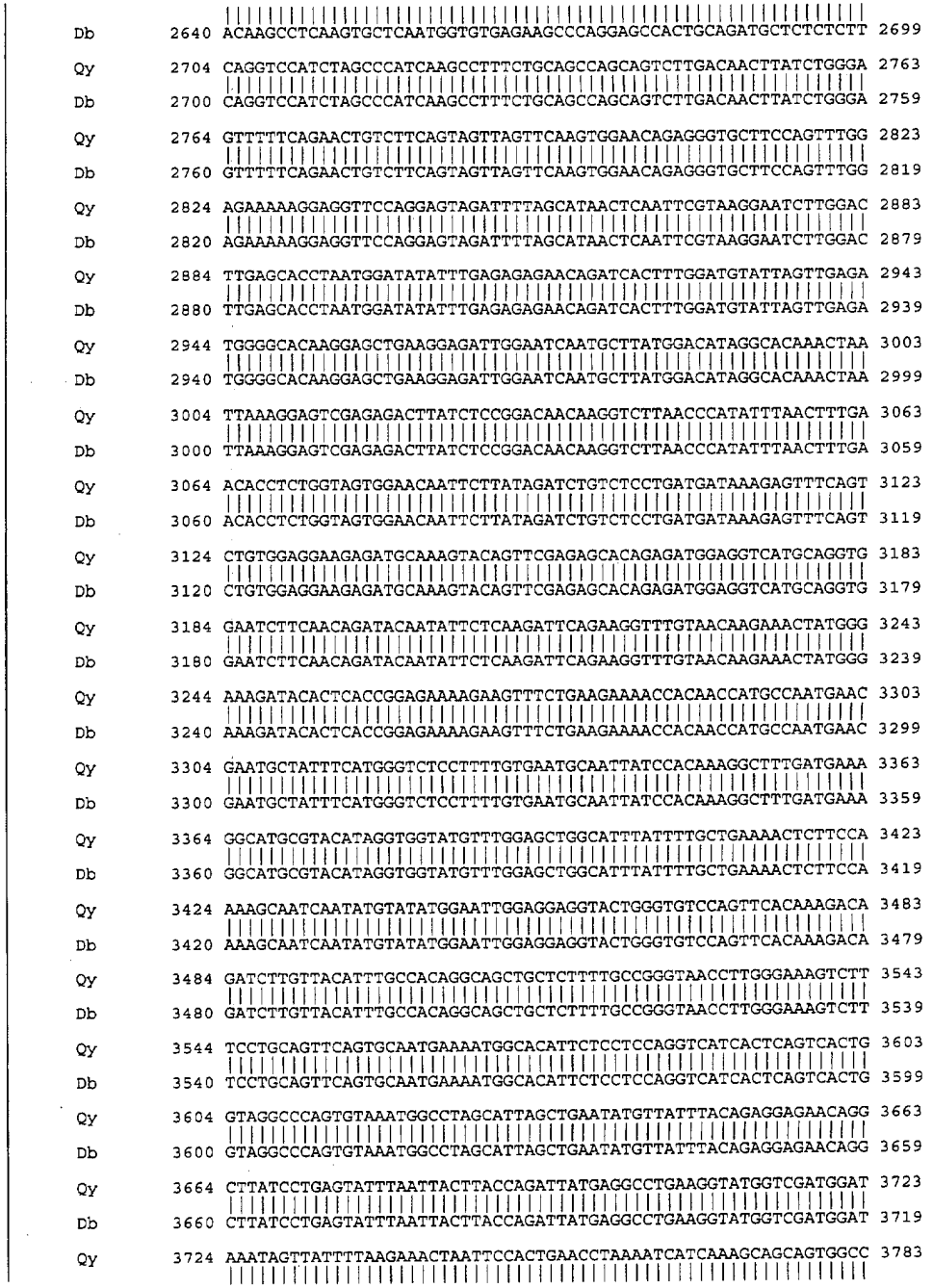
Qy       1324 CATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGC 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 CATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGC 1379

Qy       1384 TAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGG 1439

Qy       1444 CATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTA 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 CATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTA 1499

Qy       1504 ATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTAC 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 ATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTAC 1559
```





```
Db       3720 AAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCC 3779

Qy       3784 TCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
              ||||||||||||||||||||||||||||||
Db       3780 TCTACGTTTTACTCCTTTGCTGAAAAAAAA 3809


RESULT 9
ABZ11674
ID   ABZ11674 standard; cDNA; 5075 BP.
XX
AC   ABZ11674;
XX
DT   20-JAN-2003  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 556.
XX
KW   Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW   cell-proliferative disorder; neurodegenerative disease; bacterial;
KW   Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW   multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW   arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW   antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW   haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
KW   antiarthritic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200270539-A2.
XX
PD   12-SEP-2002.
XX
PF   05-MAR-2002; 2002WO-US005095.
XX
PR   05-MAR-2001; 2001US-00799451.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Zhou P,  Goodrich RW,  Asundi V,  Zhang J,  Zhao QA,  Ren F;
PI   Xue AJ,  Yang Y,  Ma Y,  Yamazaki V,  Chen R,  Wang Z,  Ghosh M;
PI   Wehrman T,  Wang J,  Wang D,  Drmanac RT;
XX
DR   WPI; 2002-759812/82.
DR   P-PSDB; ABP69457.
XX
PT   New polynucleotides comprising sequences assembled from expressed
PT   sequence tags (ESTs), useful for treating cell-proliferative,
PT   neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT   or coagulation disorders.
XX
PS   Claim 1; SEQ ID NO 556; 1012pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising a
CC   nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC   ABZ12066) or their mature protein coding portion, active domain coding
CC   protein or complementary sequences. The polynucleotides are useful for
CC   identifying expressed genes or for physical mapping of human genome. The
CC   encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC   markers, as a food supplement, for generating antibodies, in medical
CC   imaging, screening and diagnostic assays and for treating cell-
CC   proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC   or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC   diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC   platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC   or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC   arthritis, etc. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 5075 BP; 1487 A; 1055 C; 1166 G; 1367 T; 0 U; 0 Other;

  Query Match             98.8%;  Score 3770.8;  DB 6;  Length 5075;
  Best Local Similarity   99.6%;  Pred. No. 0;

  Matches 3800;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    2;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||  |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 88

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 147

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCT-CGCGGGGCCGGGGCTCCTG 179
              ||||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        148 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCCCCGGGGCCGGGGCTCCTG 207

Qy        180 CTCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGC 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 CTCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGC 267

Qy        240 TGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCC 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCC 327

Qy        300 CGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACG 359
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        328 CGAGAGCTGTTCGAGGCGTGCCCCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACG 387

Qy        360 CCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCC 419
              |||||||||||||||||||||||||||||||||||||| | ||||||||||||||||||
Db        388 CCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAAATCCACCCGCTGCACTTCGCC 447

Qy        420 GCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAA 479
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        448 GCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAAAATGGTGCAAATGTCCAA 507

Qy        480 GCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAA 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAA 567

Qy        540 GTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTAT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTAT 627

Qy        600 ACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAG 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 ACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAG 687

Qy        660 CATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGAT 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 CATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGAT 747

Qy        720 CCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCC 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCC 807

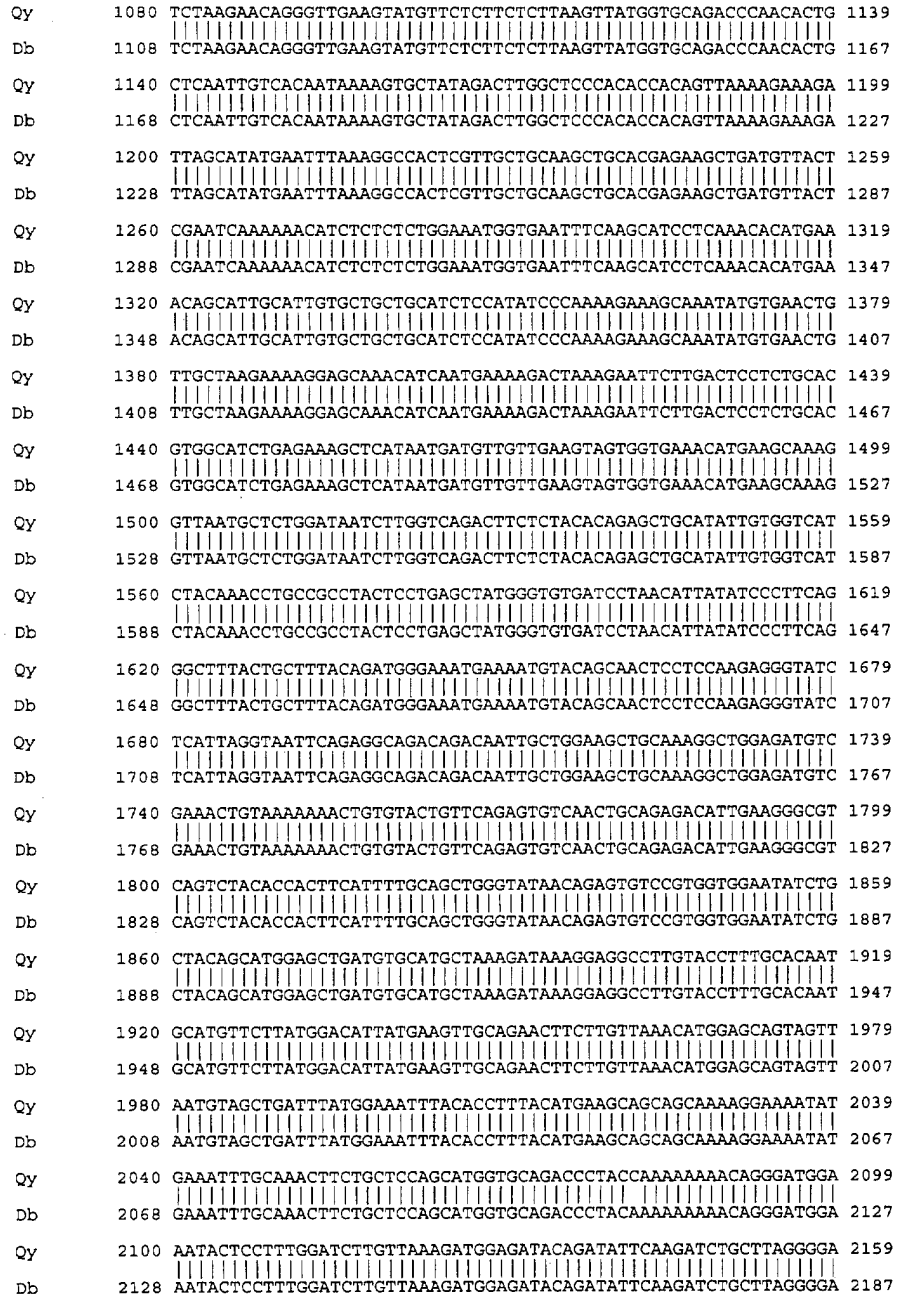
Qy        780 AGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCAC 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCAC 867

Qy        840 GCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAG 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 GCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAG 927

Qy        900 ATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTG 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 ATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTG 987

Qy        960 GTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAG 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 GTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAG 1047

Qy       1020 CATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCT 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 CATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCT 1107
```


Qy 1080 TCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTG 1139
Db 1108 TCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTG 1167
Qy 1140 CTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGA 1199
Db 1168 CTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGA 1227
Qy 1200 TTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACT 1259
Db 1228 TTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACT 1287
Qy 1260 CGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAA 1319
Db 1288 CGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAA 1347
Qy 1320 ACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTG 1379
Db 1348 ACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTG 1407
Qy 1380 TTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCAC 1439
Db 1408 TTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCAC 1467
Qy 1440 GTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAG 1499
Db 1468 GTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAG 1527
Qy 1500 GTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCAT 1559
Db 1528 GTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCAT 1587
Qy 1560 CTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAG 1619
Db 1588 CTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAG 1647
Qy 1620 GGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATC 1679
Db 1648 GGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATC 1707
Qy 1680 TCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTC 1739
Db 1708 TCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTC 1767
Qy 1740 GAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGT 1799
Db 1768 GAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGT 1827
Qy 1800 CAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTG 1859
Db 1828 CAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTG 1887
Qy 1860 CTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAAT 1919
Db 1888 CTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAAT 1947
Qy 1920 GCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTT 1979
Db 1948 GCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTT 2007
Qy 1980 AATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATAT 2039
Db 2008 AATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATAT 2067
Qy 2040 GAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGA 2099
Db 2068 GAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGA 2127
Qy 2100 AATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGA 2159
Db 2128 AATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGA 2187
Qy 2160 GATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCT 2219
Db 2188 GATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCT 2247
Qy 2220 TCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTA 2279
Db 2248 TCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTA 2307
Qy 2280 GCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTG 2339
Db 2308 GCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTG 2367
Qy 2340 AATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTA 2399
Db 2368 AATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTA 2427
Qy 2400 GATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCT 2459
Db 2428 GATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCT 2487
Qy 2460 TTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTA 2519
Db 2488 TTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTA 2547
Qy 2520 GCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTT 2579
Db 2548 GCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTT 2607
Qy 2580 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2639
Db 2608 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2667
Qy 2640 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2699
Db 2668 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2727
Qy 2700 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2759
Db 2728 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2787
Qy 2760 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2819
Db 2788 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2847
Qy 2820 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2879
Db 2848 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2907
Qy 2880 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2939
Db 2908 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2967
Qy 2940 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 2999
Db 2968 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 3027
Qy 3000 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3059
Db 3028 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3087
Qy 3060 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3119
Db 3088 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3147
Qy 3120 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3179
Db 3148 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3207
Qy 3180 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3239
Db 3208 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3267
Qy 3240 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3299

```
Db       3268 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3327

Qy       3300 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3359

Db       3328 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3387

Qy       3360 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3419

Db       3388 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3447

Qy       3420 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3479

Db       3448 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3507

Qy       3480 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3539

Db       3508 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3567

Qy       3540 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3599

Db       3568 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3627

Qy       3600 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAA 3659

Db       3628 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAA 3687

Qy       3660 CAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGAT 3719

Db       3688 CAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGAT 3747

Qy       3720 GGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGT 3779

Db       3748 GGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGT 3807

Qy       3780 GGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813

Db       3808 GGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3841
```

RESULT 10
AAF63926
ID AAF63926 standard; DNA; 4127 BP.
XX
AC AAF63926;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 100.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66288.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 158-162; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 4127 BP; 1245 A; 813 C; 934 G; 1135 T; 0 U; 0 Other;

Query Match 94.2%; Score 3595.8; DB 5; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274

Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334

Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394

Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454

Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514

Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574

Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634

Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694

Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754

Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814

Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600

Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874

Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934

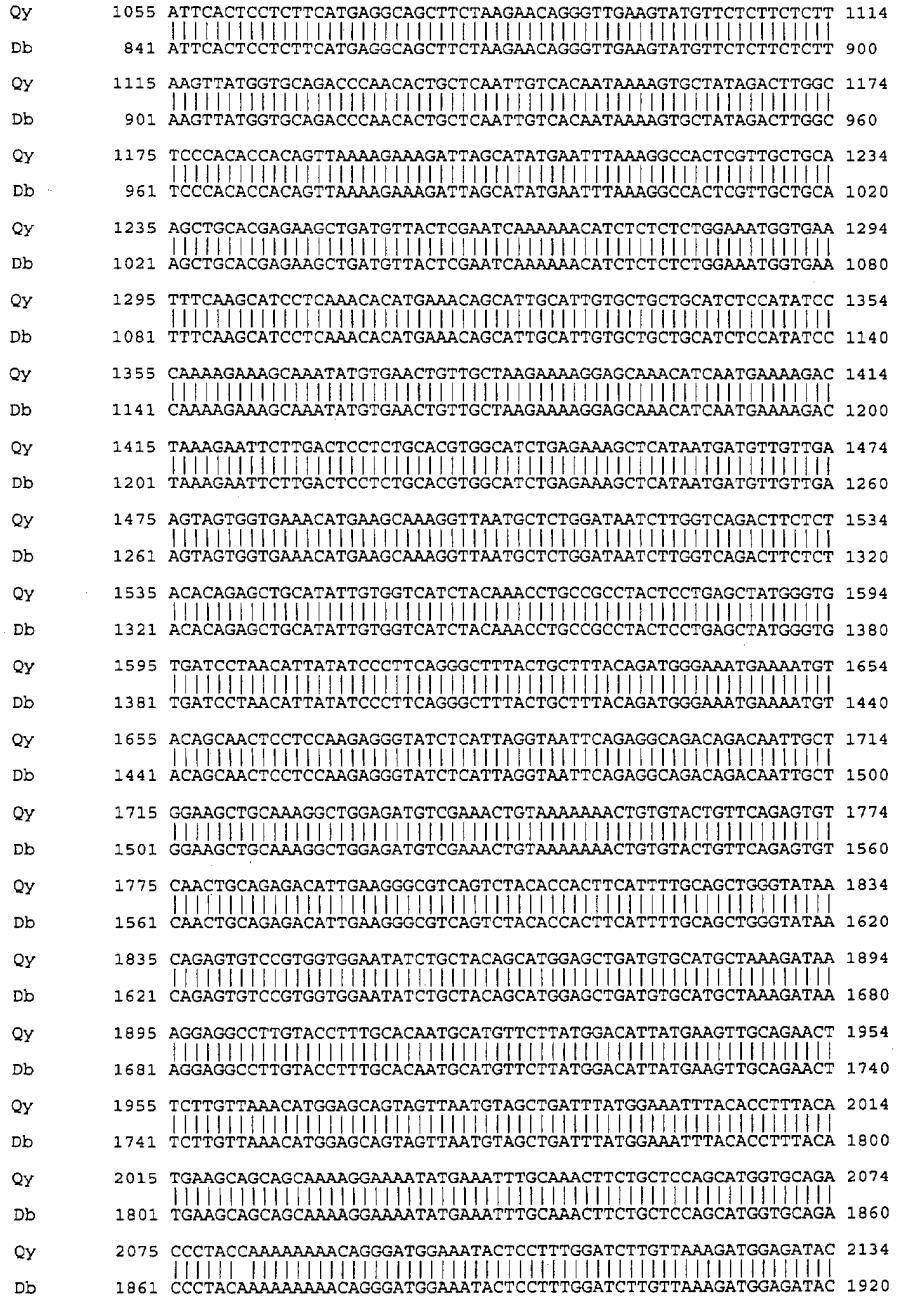
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

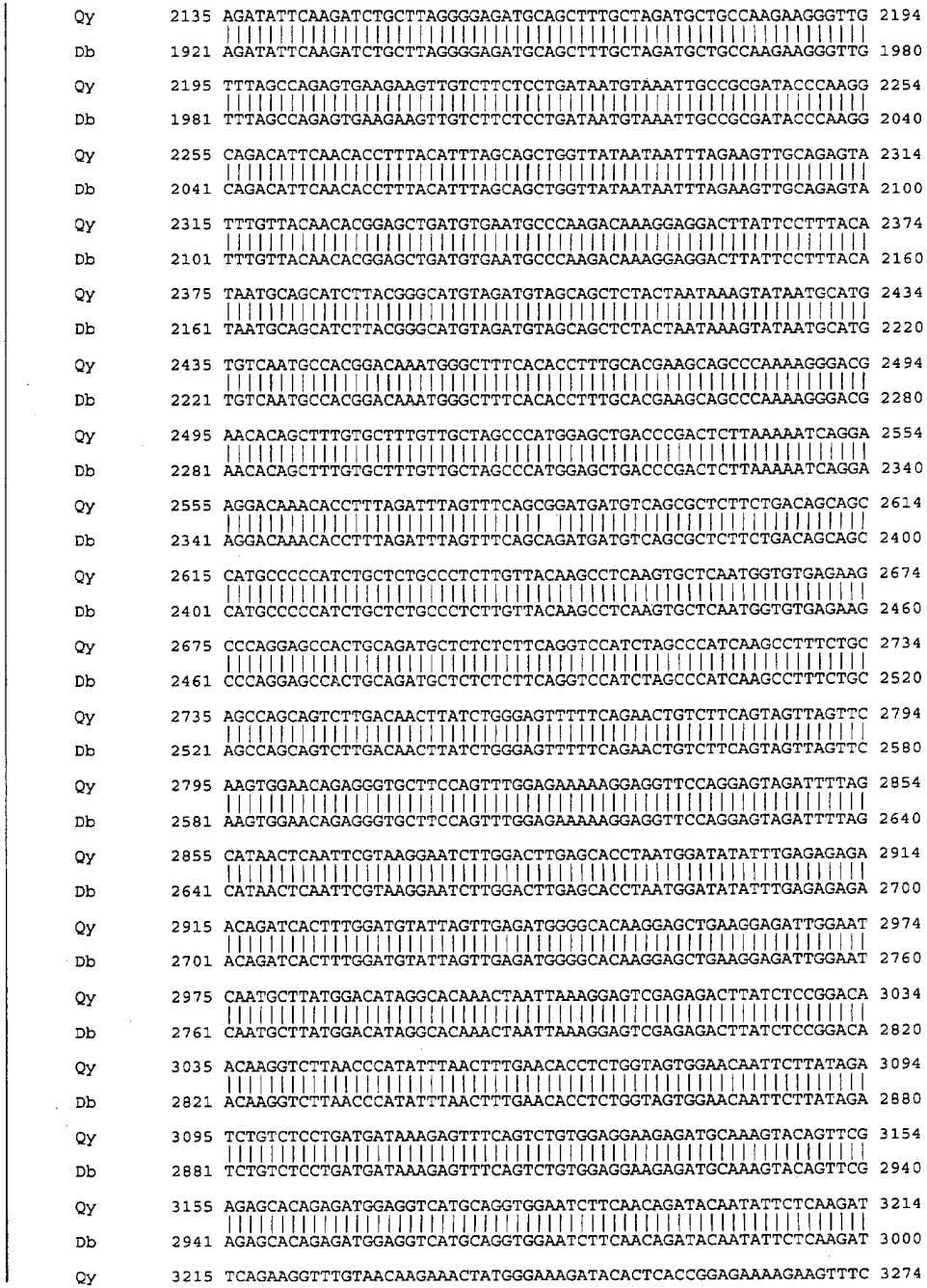
Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994

Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054

Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840
```


Qy 1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
Db 841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900
Qy 1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
Db 901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960
Qy 1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
Db 961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020
Qy 1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
Db 1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080
Qy 1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
Db 1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140
Qy 1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
Db 1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200
Qy 1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
Db 1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260
Qy 1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
Db 1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320
Qy 1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
Db 1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380
Qy 1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
Db 1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440
Qy 1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
Db 1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500
Qy 1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGT 1774
Db 1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGT 1560
Qy 1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1834
Db 1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1620
Qy 1835 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894
Db 1621 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680
Qy 1895 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1954
Db 1681 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1740
Qy 1955 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 2014
Db 1741 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 1800
Qy 2015 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
Db 1801 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860
Qy 2075 CCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 2134
Db 1861 CCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 1920


Qy 2135 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
Db 1921 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980
Qy 2195 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2254
Db 1981 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2040
Qy 2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2314
Db 2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2100
Qy 2315 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
Db 2101 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160
Qy 2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2434
Db 2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2220
Qy 2435 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
Db 2221 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280
Qy 2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
Db 2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340
Qy 2555 AGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
Db 2341 AGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGC 2400
Qy 2615 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2674
Db 2401 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2460
Qy 2675 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2734
Db 2461 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2520
Qy 2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
Db 2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580
Qy 2795 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2854
Db 2581 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2640
Qy 2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2914
Db 2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2700
Qy 2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2974
Db 2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2760
Qy 2975 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 3034
Db 2761 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 2820
Qy 3035 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 3094
Db 2821 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 2880
Qy 3095 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 3154
Db 2881 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 2940
Qy 3155 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3214
Db 2941 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3000
Qy 3215 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3274

```
Db       3001 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3060

Qy       3275 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3334

Db       3061 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3120

Qy       3335 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3394

Db       3121 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3180

Qy       3395 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3454

Db       3181 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3240

Qy       3455 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3514

Db       3241 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3300

Qy       3515 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3574

Db       3301 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3360

Qy       3575 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3634

Db       3361 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3420

Qy       3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3694

Db       3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3480

Qy       3695 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3754

Db       3481 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3540

Qy       3755 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813

Db       3541 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3599
```

RESULT 11
AAD02578

```
ID   AAD02578 standard; DNA; 4512 BP.
XX
AC   AAD02578;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human tankyrase homolog protein (THP) DNA.
XX
KW   Human; tankyrase homolog protein; THP; gene therapy; cancer; tumour;
KW   basal cell carcinoma; therapy; genetic mapping; cytostatic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             23. .3523
FT                   /*tag=  a
FT                   /product= "Human tankyrase homolog protein (THP)"
FT                   /note= "This region is specifically claimed in claim 4 as
FT                   SEQ ID NO:4"
XX
PN   WO200104326-A1.
XX
PD   18-JAN-2001.
XX
PF   03-JUL-2000; 2000WO-EP006609.
XX
PR   09-JUL-1999;   99US-00350982.
XX
PA   (PHAA ) PHARMACIA & UPJOHN SPA.
XX
PI   Berthelsen J,  Toma S,  Isacchi A;
XX
DR   WPI; 2001-168422/17.
DR   P-PSDB; AAY72589.
XX
PT   New tankyrase homolog protein (THP) polynucleotide and polypeptide useful
PT   in gene therapy, diagnosis and treatment or prevention of unregulated
PT   cell growth, such as cancer or tumor cell growth.
XX
PS   Claim 1; Page 41-42; 60pp; English.
XX
CC   The present sequence is a DNA coding for human tankyrase homolog protein
CC   (THP). The THP polypeptides and polynucleotides of the invention are
CC   useful in gene therapy and for treating or preventing unregulated cell
CC   growth such as cancer or tumour (e.g. basal cell carcinoma). The nucleic
CC   acid molecules of the invention and their fragments are useful for
CC   restriction fragment length polymorphism (RFLP) associated with certain
CC   disorders, as well as for genetic mapping. Antisense oligonucleotides, or
CC   fragments of nucleic acid encoding THP are useful as diagnostic tools for
CC   probing the expression of Thp gene in various tissues. THP can be used as
CC   antigens for raising antibodies against them and in assays for
CC   identifying compounds that modulate their activity. They are used in the
CC   manufacture of a medicament directed towards cancers or tumours. THP are
CC   also useful for screening compounds in a variety of drugs screening
CC   techniques and as a research tool for identification, characterisation
CC   and purification of interacting, regulatory proteins
XX
SQ   Sequence 4512 BP; 1356 A; 878 C; 1002 G; 1268 T; 0 U; 8 Other;
```

Query Match 94.2%; Score 3594.8; DB 4; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3593; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```
Qy        210 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 269

Db          8 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 67

Qy        270 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 329

Db         68 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 127

Qy        330 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 389

Db        128 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 187

Qy        390 GGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 449

Db        188 GGCAGGAAATCCACCCCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 247

Qy        450 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 509

Db        248 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 307

Qy        510 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 569

Db        308 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 367

Qy        570 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 629

Db        368 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 427

Qy        630 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 689

Db        428 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 487

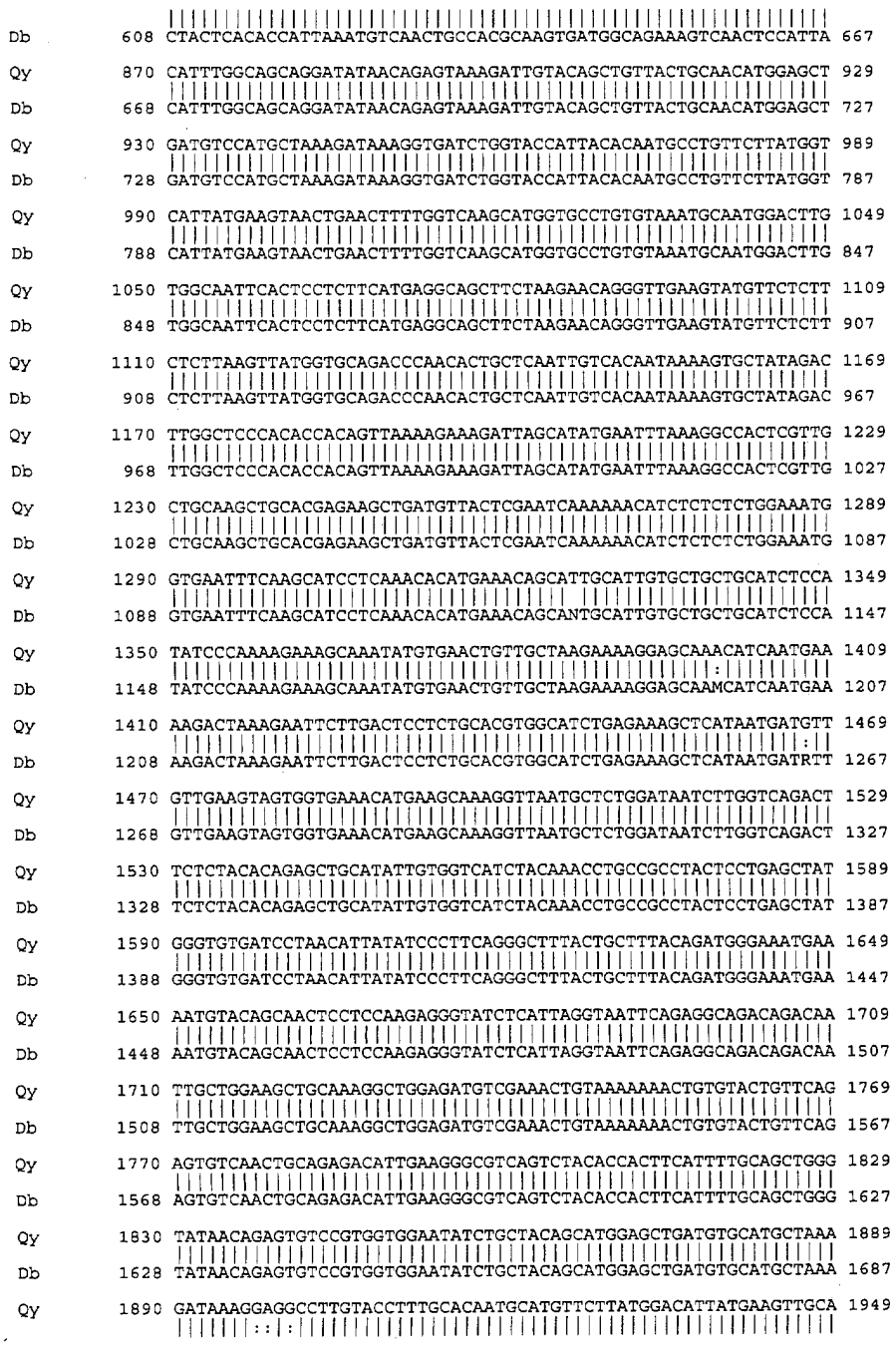
Qy        690 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 749

Db        488 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 547

Qy        750 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 809

Db        548 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 607

Qy        810 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 869
```


Db 608 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 667
Qy 870 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 929
Db 668 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 727
Qy 930 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 989
Db 728 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 787
Qy 990 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 1049
Db 788 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 847
Qy 1050 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 1109
Db 848 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 907
Qy 1110 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 1169
Db 908 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 967
Qy 1170 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229
Db 968 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027
Qy 1230 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1289
Db 1028 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1087
Qy 1290 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCA 1349
Db 1088 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCA 1147
Qy 1350 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAA 1409
Db 1148 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAA 1207
Qy 1410 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTT 1469
Db 1208 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTT 1267
Qy 1470 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1529
Db 1268 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1327
Qy 1530 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1589
Db 1328 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1387
Qy 1590 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1649
Db 1388 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1447
Qy 1650 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1709
Db 1448 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1507
Qy 1710 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1769
Db 1508 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1567
Qy 1770 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1829
Db 1568 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1627
Qy 1830 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1889
Db 1628 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1687
Qy 1890 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1949
Db 1688 GATAAAGRRGSCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1747
Qy 1950 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 2009
Db 1748 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 1807
Qy 2010 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069
Db 1808 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 1867
Qy 2070 GCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 2129
Db 1868 GCAGACCCTACAAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 1927
Qy 2130 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 2189
Db 1928 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 1987
Qy 2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
Db 1988 AGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2047
Qy 2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
Db 2048 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2107
Qy 2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
Db 2108 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2167
Qy 2370 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2429
Db 2168 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2227
Qy 2430 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2489
Db 2228 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2287
Qy 2490 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2549
Db 2288 GGACGAACACAGCTTTGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2347
Qy 2550 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2609
Db 2348 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2407
Qy 2610 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2669
Db 2408 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2467
Qy 2670 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729
Db 2468 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2527
Qy 2730 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2789
Db 2528 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2587
Qy 2790 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2849
Db 2588 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2647
Qy 2850 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2909
Db 2648 TTTAGCATAACTCAATTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2707
Qy 2910 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2969
Db 2708 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2767
Qy 2970 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 3029
Db 2768 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 2827

```
Qy       3030 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 3089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2828 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 2887

Qy       3090 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 3149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2888 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 2947

Qy       3150 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2948 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3007

Qy       3210 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3008 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3067

Qy       3270 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3068 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3127

Qy       3330 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3128 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3187

Qy       3390 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3247

Qy       3450 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3248 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3307

Qy       3510 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3308 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3367

Qy       3570 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3368 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3427

Qy       3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3428 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3487

Qy       3690 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3488 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3547

Qy       3750 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3548 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3607

Qy       3810 AAAA 3813
             ||||
Db       3608 AAAA 3611
```

```
RESULT 12
AAF63837
ID   AAF63837 standard; DNA; 3508 BP.
XX
AC   AAF63837;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 related coding sequence SEQ ID NO: 1.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66278.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Disclosure; Page 105-109; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 3508 BP; 1046 A; 710 C; 843 G; 909 T; 0 U; 0 Other;
```

Query Match 91.8%; Score 3504.8; DB 5; Length 3508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3506; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

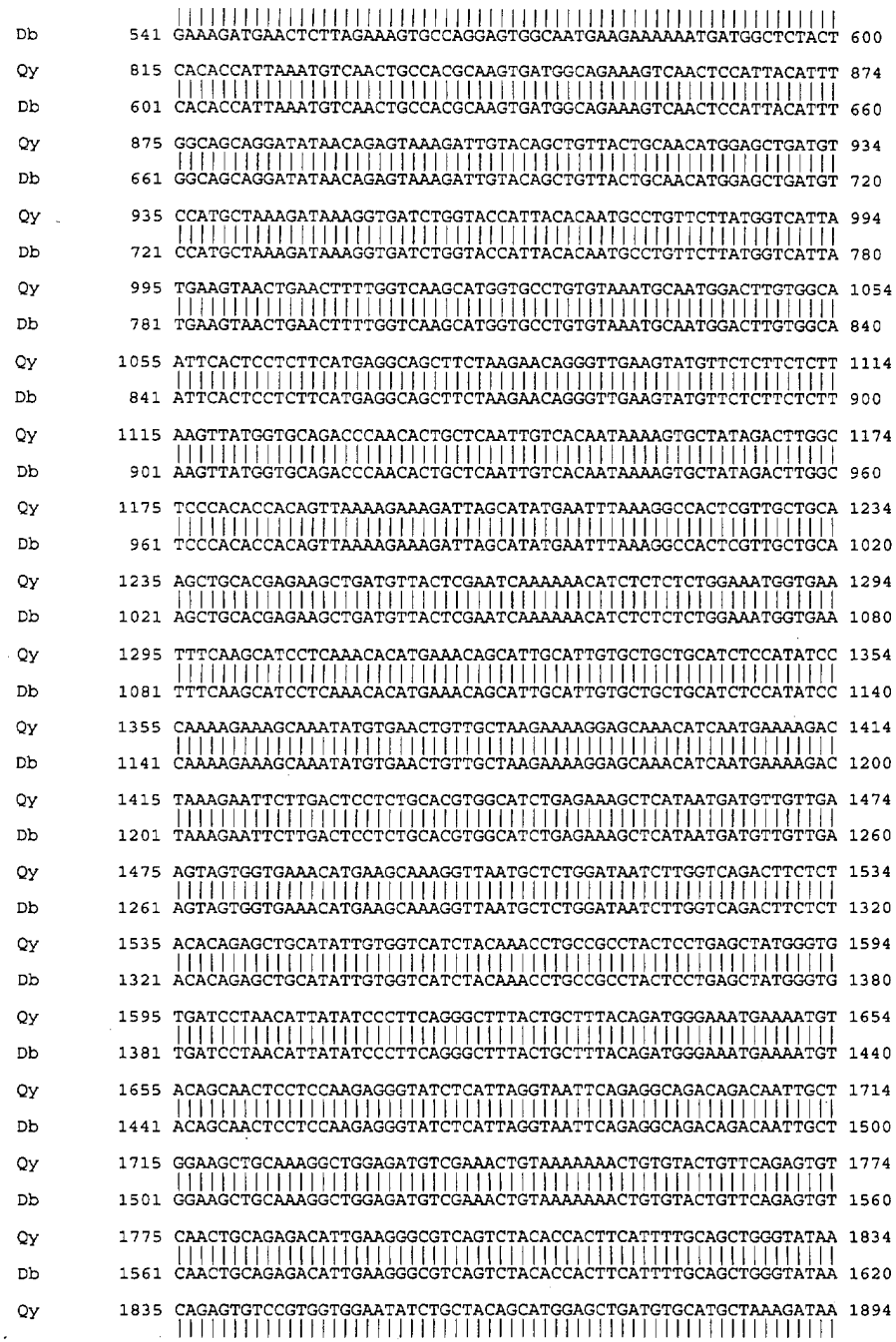
Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
```


Db 541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600
Qy 815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
Db 601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660
Qy 875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
Db 661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720
Qy 935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
Db 721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780
Qy 995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
Db 781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840
Qy 1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
Db 841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900
Qy 1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
Db 901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960
Qy 1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
Db 961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020
Qy 1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
Db 1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080
Qy 1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
Db 1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140
Qy 1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
Db 1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200
Qy 1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
Db 1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260
Qy 1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
Db 1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320
Qy 1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
Db 1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380
Qy 1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
Db 1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440
Qy 1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
Db 1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500
Qy 1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
Db 1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560
Qy 1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1834
Db 1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1620
Qy 1835 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894
Db 1621 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680
Qy 1895 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1954
Db 1681 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1740
Qy 1955 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 2014
Db 1741 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 1800
Qy 2015 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
Db 1801 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860
Qy 2075 CCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 2134
Db 1861 CCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 1920
Qy 2135 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
Db 1921 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980
Qy 2195 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2254
Db 1981 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2040
Qy 2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2314
Db 2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2100
Qy 2315 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
Db 2101 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160
Qy 2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2434
Db 2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2220
Qy 2435 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
Db 2221 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280
Qy 2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
Db 2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340
Qy 2555 AGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
Db 2341 AGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGC 2400
Qy 2615 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2674
Db 2401 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2460
Qy 2675 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2734
Db 2461 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2520
Qy 2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
Db 2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580
Qy 2795 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2854
Db 2581 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2640
Qy 2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2914
Db 2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2700
Qy 2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2974
Db 2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2760

```
Qy       2975 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 3034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 2820

Qy       3035 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 3094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 2880

Qy       3095 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 3154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 2940

Qy       3155 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3000

Qy       3215 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3060

Qy       3275 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3120

Qy       3335 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3180

Qy       3395 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3240

Qy       3455 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3300

Qy       3515 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3360

Qy       3575 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3420

Qy       3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3480

Qy       3695 TATGAGGCCTGAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||||||||||||
Db       3481 TATGAGGCCTGAAGGTATGGTCGATGGA 3508
```

```
RESULT 13
AAA91487
ID   AAA91487 standard; DNA; 3797 BP.
XX
AC   AAA91487;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Tankyrase homologue isotype 1 coding sequence.
XX
KW   Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW   cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW   chromosome mapping; gene therapy; vaccine; ds.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   CDS             404. .3706
FT                   /*tag=  a
FT                   /partial
FT                   /product= "TaHo1"
FT                   /note= "Tankyrase homologue isotype 1; No start codon
FT                   given"
XX
PN   WO200130987-A2.
XX
PD   03-MAY-2001.
XX
PF   25-OCT-2000; 2000WO-US041528.
XX
PR   25-OCT-1999;   99US-00427154.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B;
XX
DR   WPI; 2001-300503/31.
DR   P-PSDB; AAY97748.
XX
PT   Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT   or preventing cell proliferation in cells, and for diagnosing, treating
PT   or preventing cell cycle associated disorders such as cancer.
XX
PS   Claim 13; Fig 1; 63pp; English.
XX
CC   This sequence encodes the Tankyrase homologue isotype 1 (TaHo-1) protein
CC   of the invention. The invention also relates to the TaHo-2 protein. The
CC   TaHo proteins are useful for inducing or preventing cell proliferation in
CC   cells, and in the study or treatment of conditions mediated by the cell
CC   cycle proteins, such as to diagnose, treat or prevent cell cycle
CC   associated disorders, preferably cancer. The TaHo coding sequences are
CC   useful as hybridisation probes, in chromosome and gene mapping and in the
CC   generation of anti-sense DNA and RNA. The coding sequences are also
CC   useful for the preparation of TaHo, for generating either transgenic
CC   animals or knock out animals which, in turn, are useful in a development
CC   and screening of therapeutically useful agents, in gene therapy, as
CC   vaccine, and for construction of hybridisation probes for mapping the
CC   gene which encodes TaHo and for the genetic analysis of individuals with
CC   genetic disorders. The TaHo proteins, and their coding sequences are
CC   useful in screening assays
XX
SQ   Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 U; 0 Other;

  Query Match             88.9%;  Score 3393.4;  DB 4;  Length 3797;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3394;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Qy        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522

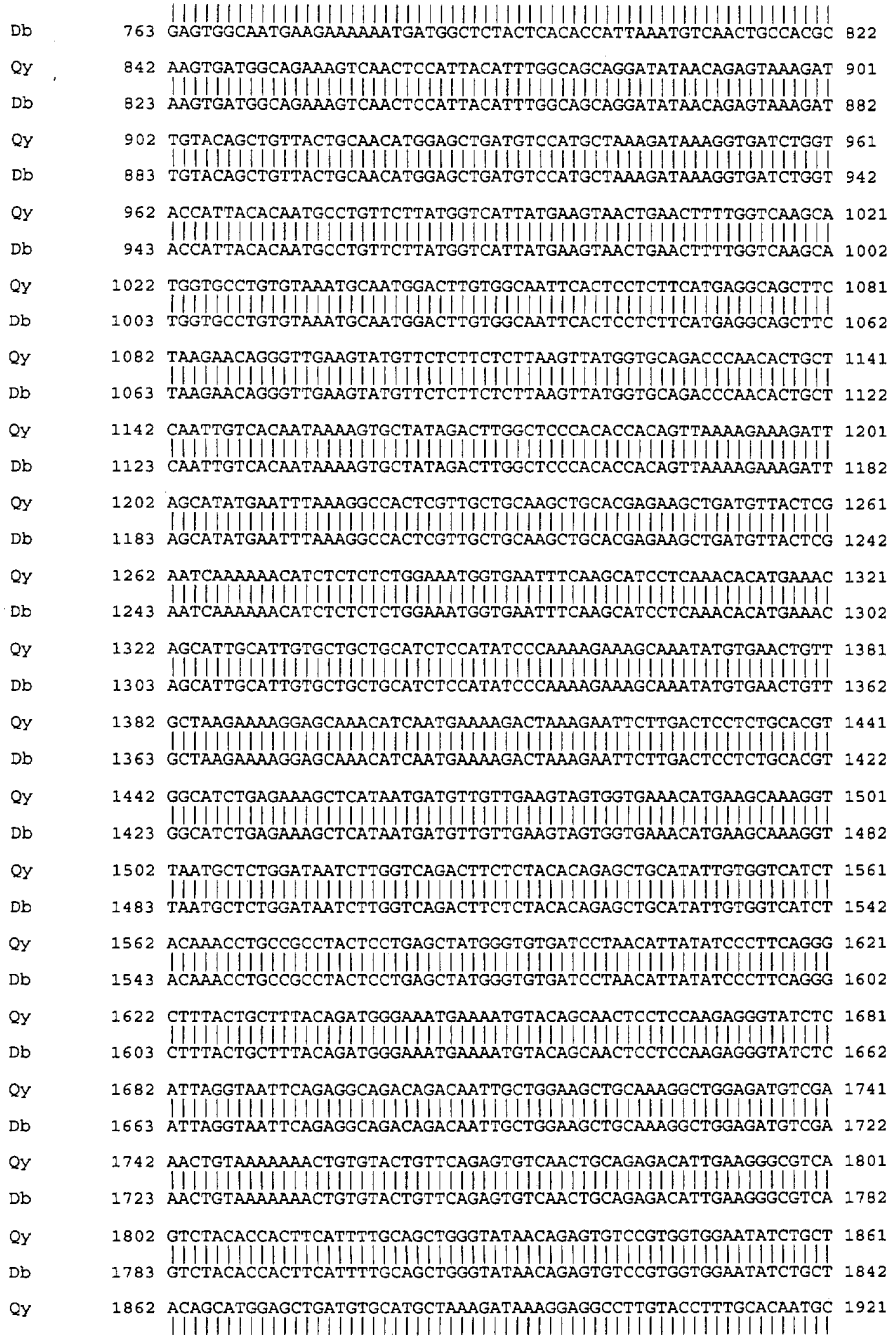
Qy        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582

Qy        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642

Qy        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702

Qy        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762

Qy        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
```


Db 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Qy 842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
Db 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Qy 902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
Db 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Qy 962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
Db 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Qy 1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
Db 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Qy 1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
Db 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Qy 1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
Db 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Qy 1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
Db 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Qy 1262 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321
Db 1243 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Qy 1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381
Db 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Qy 1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Db 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Qy 1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Db 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Qy 1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Db 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Qy 1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Db 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Qy 1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Db 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Qy 1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Db 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Qy 1742 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Db 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Qy 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Db 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Qy 1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Db 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Qy 1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Db 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Qy 1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Db 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Qy 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2101
Db 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Qy 2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Db 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Qy 2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Qy 2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Qy 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Qy 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Qy 2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Qy 2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Qy 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Qy 2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Db 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Qy 2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Db 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Qy 2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Db 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Qy 2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Db 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Qy 2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Db 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Qy 2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Db 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Qy 2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Db 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982

```
Qy       3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042

Qy       3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102

Qy       3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162

Qy       3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

Qy       3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

Qy       3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342

Qy       3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402

Qy       3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462

Qy       3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

Qy       3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Qy       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Qy       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Qy       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Qy       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
             ||||||||||||||||||||||||||||||||||||
Db       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
```

RESULT 14
ABQ77066
ID ABQ77066 standard; DNA; 3797 BP.
XX
AC ABQ77066;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 1 TaHo-1 DNA from clone TH-1.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-1;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 404. .3706
FT /*tag= a
FT /product= "TaHo-1"
FT /partial
FT /note= "No start codon given"
XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
DR P-PSDB; ABG73730.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Claim 13; Fig 1A-B; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence encodes the tankyrase H isoform 1, TaHo-1, isolated from clone
CC TH-1 which is described in the disclosure of the invention
XX
SQ Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 U; 0 Other;

Query Match 88.9%; Score 3393.4; DB 7; Length 3797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Qy        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522

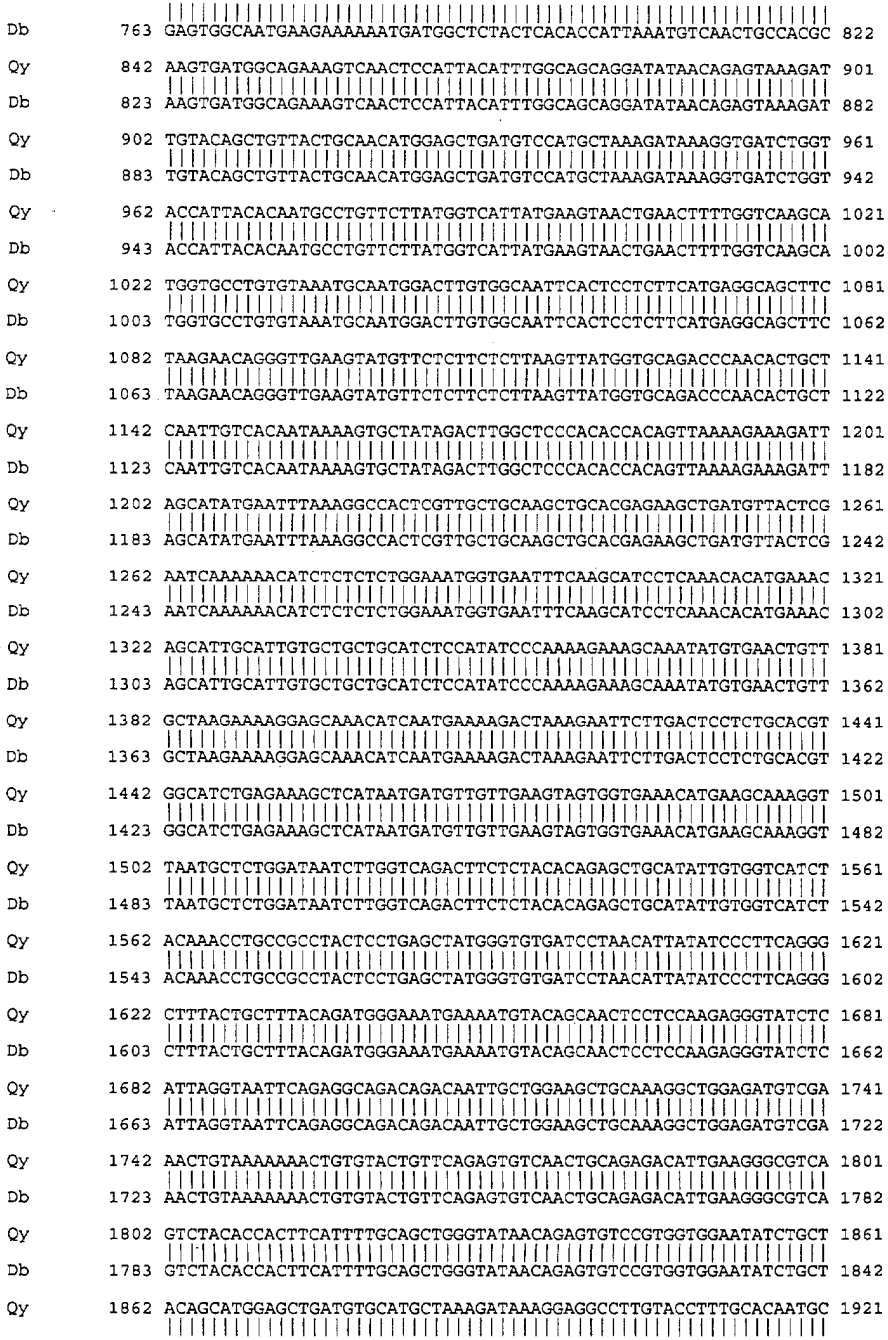
Qy        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582

Qy        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642

Qy        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702

Qy        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762

Qy        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
```


Db 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Qy 842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
Db 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Qy 902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
Db 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Qy 962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
Db 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Qy 1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
Db 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Qy 1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
Db 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Qy 1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
Db 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Qy 1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
Db 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Qy 1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321
Db 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Qy 1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381
Db 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Qy 1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Db 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Qy 1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Db 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Qy 1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Db 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Qy 1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Db 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Qy 1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Db 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Qy 1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Db 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Qy 1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Db 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Qy 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Db 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Qy 1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Db 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Qy 1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Db 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Qy 1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Db 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Qy 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2101
Db 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Qy 2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Db 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Qy 2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Qy 2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Qy 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Qy 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Qy 2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Qy 2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Qy 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Qy 2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Db 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Qy 2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Db 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Qy 2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Db 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Qy 2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Db 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Qy 2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Db 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Qy 2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Db 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Qy 2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Db 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982

```
Qy       3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042

Qy       3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102

Qy       3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162

Qy       3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

Qy       3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

Qy       3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342

Qy       3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402

Qy       3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462

Qy       3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

Qy       3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Qy       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Qy       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Qy       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Qy       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
             ||||||||||||||||||||||||||||||||||||
Db       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797


RESULT 15
AAC66824
ID   AAC66824 standard; cDNA; 4296 BP.
XX
AC   AAC66824;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II coding sequence SEQ ID NO: 3.
XX
KW   Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .4373
FT                   /*tag=  a
FT                   /transl_except= (pos:182. .184,aa:Xaa)
FT                   /transl_except= (pos:350. .352,aa:Xaa)
FT                   /transl_except= (pos:385. .387,aa:Xaa)
FT                   /transl_except= (pos:397. .399,aa:Xaa)
FT                   /transl_except= (pos:434. .436,aa:Xaa)
FT                   /transl_except= (pos:473. .475,aa:Xaa)
FT                   /transl_except= (pos:554. .556,aa:Xaa)
FT                   /transl_except= (pos:575. .580,aa:XaaXaa)
FT                   /transl_except= (pos:590. .592,aa:Xaa)
FT                   /transl_except= (pos:611. .613,aa:Xaa)
FT                   /transl_except= (pos:620. .622,aa:Xaa)
FT                   /transl_except= (pos:662. .664,aa:Xaa)
FT                   /transl_except= (pos:668. .673,aa:XaaXaa)
FT                   /transl_except= (pos:677. .679,aa:Xaa)
FT                   /transl_except= (pos:689. .691,aa:Xaa)
FT                   /transl_except= (pos:95. .697,aa:Xaa)
FT                   /transl_except= (pos:701. .703,aa:Xaa)
FT                   /transl_except= (pos:737. .742,aa:XaaXaa)
FT                   /transl_except= (pos:767. .772,aa:XaaXaa)
FT                   /transl_except= (pos:2057. .2059,aa:Xaa)
FT                   /transl_except= (pos:2231. .2233,aa:Xaa)
FT                   /transl_except= (pos:2237. .2239,aa:Xaa)
FT                   /transl_except= (pos:2261. .2263,aa:Xaa)
FT                   /transl_except= (pos:2279. .2281,aa:Xaa)
FT                   /transl_except= (pos:2345. .2350,aa:XaaXaa)
FT                   /transl_except= (pos:3035. .3037,aa:Xaa)
FT                   /transl_except= (pos:3860. .3862,aa:Xaa)
FT                   /transl_except= (pos:3881. .3883,aa:Xaa)
FT                   /transl_except= (pos:3890. .3892,aa:Xaa)
FT                   /transl_except= (pos:3896. .3898,aa:Xaa)
FT                   /transl_except= (pos:3941. .3943,aa:Xaa)
FT                   /transl_except= (pos:4001. .4003,aa:Xaa)
FT                   /transl_except= (pos:4025. .30,aa:XaaXaa)
FT                   /transl_except= (pos:4118. .4120,aa:Xaa)
FT                   /transl_except= (pos:4133. .35,aa:Xaa)
FT                   /transl_except= (pos:4172. .4174,aa:Xaa)
FT                   /transl_except= (pos:4190. .4192,aa:Xaa)
FT                   /transl_except= (pos:4199. .4201,aa:Xaa)
FT                   /transl_except= (pos:4238. .4240,aa:Xaa)
FT                   /transl_except= (pos:4256. .4258,aa:Xaa)
FT                   /transl_except= (pos:4268. .4270,aa:Xaa)
FT                   /transl_except= (pos:4283. .4285,aa:Xaa)
FT                   /transl_except= (pos:4316. .4318,aa:Xaa)
FT                   /note= "Xaa=unknown. The sequence contains a number of in
FT                   -frame termination codons, given as Xaa in the protein
FT                   sequence"
FT                   /product= "tankyrase II"
FT                   /partial
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   P-PSDB; AAB27210.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Disclosure; Fig 3; 52pp; English.
XX
```

```
CC   The present sequence is a version of the human tankyrase II coding
CC   sequence. Its protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 4296 BP; 1161 A; 820 C; 925 G; 1033 T; 0 U; 357 Other;

  Query Match            87.6%;  Score 3341;  DB 3;  Length 4296;
  Best Local Similarity  89.0%;  Pred. No. 0;
  Matches 3400;  Conservative    8;  Mismatches  402;  Indels   12;  Gaps    2;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||  |  |||||||||||||||||||||||||||||||| |||  ||
Db         73 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAANGGGCGCGC 132

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
                ||    | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 GTGGGCGCGGCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 192

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 252

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        253 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCANGATCATGTCGGGTCGCCGCT 312

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              |||||||||||||| |||||||||||| |||||||||||| |||||||| ||||||||||
Db        313 GCGCCGGCGGGGGANCGGCCTGCGCGANCGCCGCGGCCGAAGCCGTGGAACCGGCCGCCC 372

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              || | |||||||| |||||||||||||||||||||||||| |||||| |||||||||| |
Db        373 GAAANCTGTTCGAAGCGTGCCGCAACGGGGACGTGGAACGANTCAAGAAGCTGGTGACNC 432

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||:|||||||||||||||||||| |||||||||||||||||||||||||||||: |||
Db        433 CTGARAAGGTGAACAGCCGCGACACNGCGGGCAGGAAATCCACCCCGCTGCACTTYCCCG 492

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              || |||||||||||||||| || ||:|||||||||||| |||||||||||||||:|||
Db        493 CANGTTTTGGGCGGAAAGACTTANTTRAATATTTGCTTCANAATGGTGCAAATGTYCAAN 552

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              |||||||| |||||||||||||||||||||||||||||||||||||:||||:||
Db        553 CACGTGATNATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCMTGCTRAAA 612

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
                || |||||||||||| |||| ||||| |:|||||||||||||||||||||||||||||
Db        613 NTATCNATCTCCTTTTGCNACATNGTGCANAMCCCAATGCTCGAGATAATTGGAATTATA 672

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              |||||| | |||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        673 CTCCTCNCNATGAAGCTGCAATTAAAGGAAAGATTGANNNTTGCATTGTGCTGTTACAGC 732

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 792

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 852

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 912

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        913 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 972

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1032

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1092

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1152

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1212

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1272

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1332

Qy       1261 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1392

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1393 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1452

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1453 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1512

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1513 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1572

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1573 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1632

Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1633 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1692

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1693 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1752

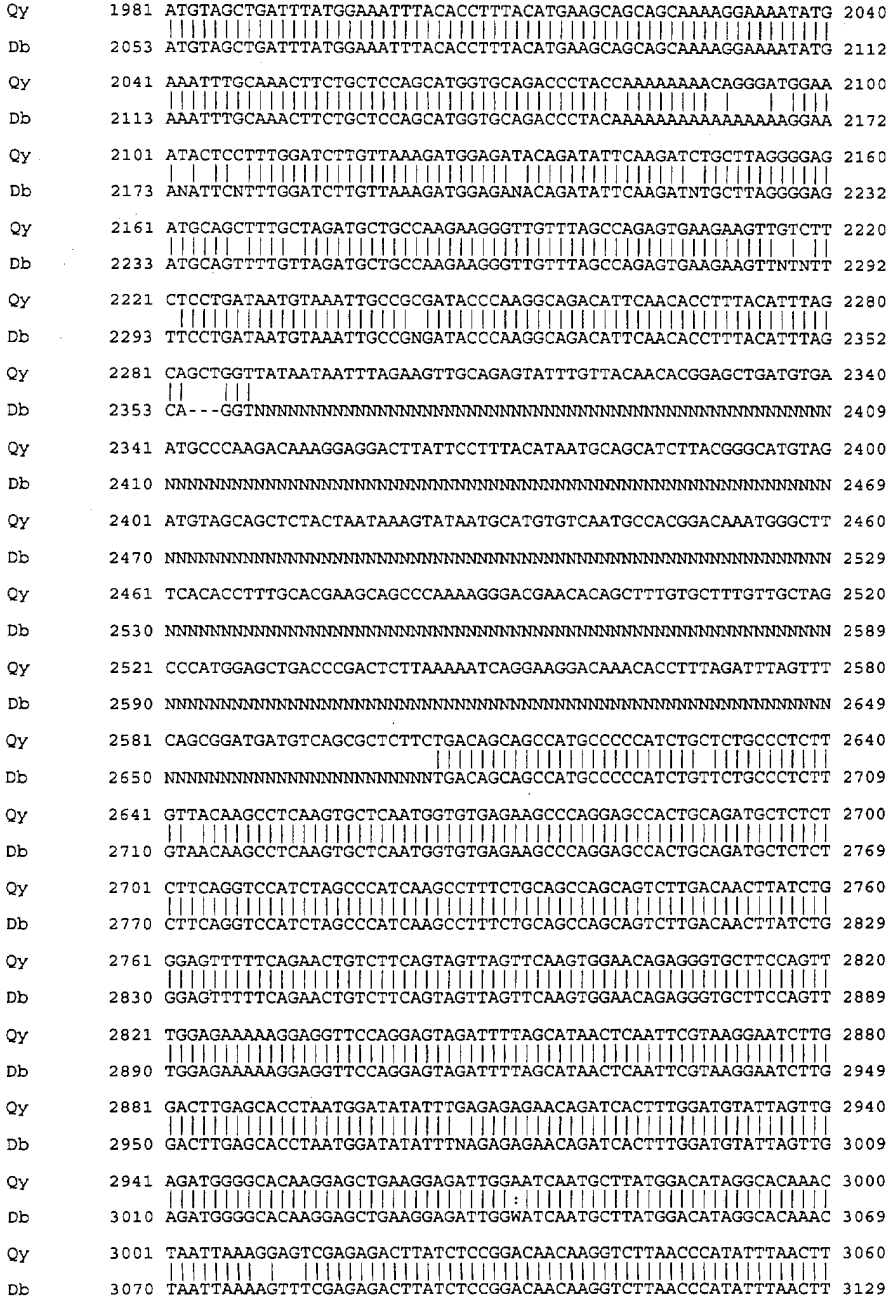
Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1753 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1812

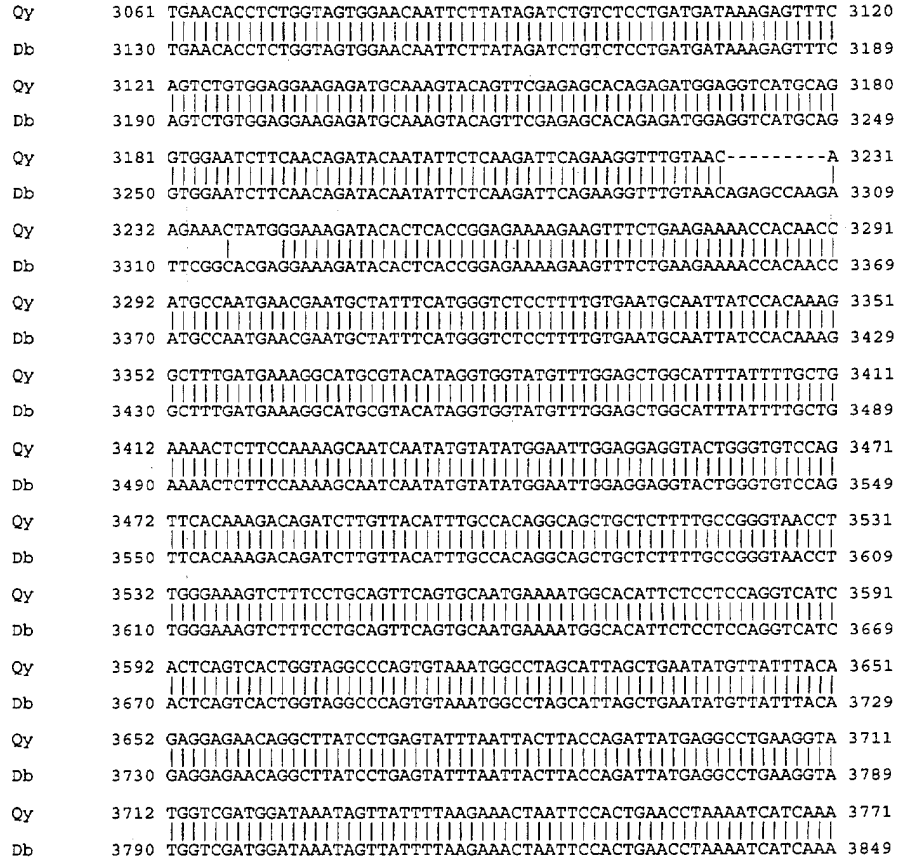
Qy       1741 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1813 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1872

Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1873 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1932

Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1933 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGNGGCCTTGTACCTTTGCACAATG 1992

Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTTAAACATGGAGCAGTAGTTA 1980
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1993 CATGTTNTTATGGACATTATGAAGTTGCAGAACTTCTTGTTTAAACATGGAGCAGTAGTTA 2052
```





RESULT 16
AAX25366
ID AAX25366 standard; cDNA; 3400 BP.
XX
AC AAX25366;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human Grb7 effector 2.2412 cDNA.
XX
KW Grb7 effector; 2.2412 protein; human; signal transduction; tumour marker;
KW breast cancer; prostate cancer; prognosis; diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .3225
FT /*tag= a
XX
PN WO9915647-A1.

```
XX
PD   01-APR-1999.
XX
PF   23-SEP-1998;   98WO-AU000795.
XX
PR   23-SEP-1997;   97AU-00009388.
XX
PA   (GARV-) GARVAN INST MEDICAL RES.
XX
PI   Daly RJ,  Sutherland RL;
XX
DR   WPI; 1999-254707/21.
DR   P-PSDB; AAY05734.
XX
PT   New candidate effector for the Grb7 family of signaling proteins, and
PT   specific antibody, useful for detection and treatment of cancer.
XX
PS   Claim 4; Fig 1; 24pp; English.
XX
CC   This is the nucleotide sequence of an isolated polynucleotide molecule
CC   encoding a candidate effector protein, termed 2.2412 (see AAY05734), for
CC   the Grb7 family of signalling proteins. A partial clone was isolated from
CC   a human liver cDNA library using a yeast two-hybrid system assay with
CC   Grb14 as bait. This clone was used to screen a human placental cDNA
CC   library that provided the C-terminal end of the 2.2412 sequence. A 5'
CC   sequence for the clone, including the initiation codon, was not
CC   identified. Analysis of the sequence revealed significant homology to a
CC   large number of proteins containing ankyrin-like repeats. The gene was
CC   localised to between chromosome 10q23.2 and proximal 10q23.32. Deletions
CC   in the 10q22-25 region have been detected in human breast, prostate,
CC   renal, small cell lung and endometrial carcinomas, glioblastoma
CC   multiforme, melanoma and meningiomas. Detection of the protein encoded by
CC   the 2.2412 cDNA in a sample should provide a useful tumour marker and/or
CC   prognostic indicator for certain human cancers, in particular breast
CC   cancer and prostate cancer. Antagonism of the interaction between Grb7
CC   family members and the encoded protein should provide a novel treatment
CC   strategy for human diseases exhibiting aberrant receptor tyrosine kinase
CC   signalling, such as cancer. Oligonucleotide probes can be used in methods
CC   of detecting the presence of 2.2412 mRNA in a sample
XX
SQ   Sequence 3400 BP; 1059 A; 665 C; 760 G; 916 T; 0 U; 0 Other;

  Query Match             86.4%;  Score 3297;  DB 2;  Length 3400;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        501 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA  60

Qy        561 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120

Qy        621 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180

Qy        681 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240

Qy        741 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300

Qy        801 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360

Qy        861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420

Qy        921 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480

Qy        981 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540

Qy       1041 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600

Qy       1101 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660

Qy       1161 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720

Qy       1221 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCT 780

Qy       1281 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840

Qy       1341 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900

Qy       1401 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960

Qy       1461 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020

Qy       1521 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080

Qy       1581 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy       1641 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200

Qy       1701 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260

Qy       1761 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320

Qy       1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380

Qy       1881 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440

Qy       1941 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500
```

```
Qy       2001 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2060
Db       1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560
Qy       2061 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2120
Db       1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620
Qy       2121 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2180
Db       1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680
Qy       2181 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2240
Db       1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740
Qy       2241 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2300
Db       1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800
Qy       2301 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2360
Db       1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860
Qy       2361 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420
Db       1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920
Qy       2421 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2480
Db       1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980
Qy       2481 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2540
Db       1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040
Qy       2541 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2600
Db       2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100
Qy       2601 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2660
Db       2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160
Qy       2661 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2720
Db       2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220
Qy       2721 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2780
Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280
Qy       2781 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2840
Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340
Qy       2841 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2900
Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400
Qy       2901 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2960
Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460
Qy       2961 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3020
Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520
Qy       3021 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3080
Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580
Qy       3081 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3140
Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640
Qy       3141 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3200
Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700
Qy       3201 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3260
Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760
Qy       3261 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3320
Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820
Qy       3321 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3380
Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880
Qy       3381 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3440
Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940
Qy       3441 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3500
Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000
Qy       3501 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3560
Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060
Qy       3561 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3620
Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120
Qy       3621 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3680
Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180
Qy       3681 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3740
Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240
Qy       3741 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3800
Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300
Qy       3801 TGCTGAAAAAAAA 3813
Db       3301 TGCTGAAAAAAAA 3313
```

RESULT 17
AAC66823
ID AAC66823 standard; cDNA; 4493 BP.
XX
AC AAC66823;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase II coding sequence SEQ ID NO: 1.
XX
KW Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS Homo sapiens.
XX

	Key	Location/Qualifiers
FH		
FT	CDS	1. .4002
FT		/*tag= a
FT		/transl_except= (pos:595. .597,aa:Xaa)
FT		/transl_except= (pos:625. .630,aa:XaaXaa)
FT		/transl_except= (pos:637. .639,aa:Xaa)
FT		/transl_except= (pos:694. .696,aa:Xaa)

```
FT                   /transl_except= (pos:724. .726,aa:Xaa)
FT                   /transl_except= (pos:730. .732,aa:Xaa)
FT                   /transl_except= (pos:763. .765,aa:Xaa)
FT                   /transl_except= (pos:790. .792,aa:Xaa)
FT                   /transl_except= (pos:2176. .2178,aa:Xaa)
FT                   /transl_except= (pos:2380. .2382,aa:Xaa)
FT                   /transl_except= (pos:2398. .2400,aa:Xaa)
FT                   /transl_except= (pos:2464. .2469,aa:XaaXaa)
FT                   /transl_except= (pos:3163. .3165,aa:Xaa)
FT                   /note= "Xaa=unknown"
FT                   /product= "tankyrase II"
FT                   /partial
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   P-PSDB; AAB27209.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Claim 1; Fig 2; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II coding
CC   sequence. Its protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 4493 BP; 1119 A; 685 C; 760 G; 985 T; 0 U; 944 Other;

  Query Match           79.6%;  Score 3039;  DB 3;  Length 4493;
  Best Local Similarity  88.5%;  Pred. No. 0;
  Matches 3101;  Conservative    1;  Mismatches  375;  Indels   27;  Gaps    2;

Qy        337 AACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGA 396
              |||  |  |   |||||||||||| ||||| |||||  |||| || | |||||| ||||||||
Db        587 AACCATTCCNNAGGCTGGTGACCCCTGAAAAGGTNAACANCCNCAACACGGNGGGCAGGA 646

Qy        397 AATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGC 456
              |||||||||||||||||||| |||||||||||||||||||| || |  || |||||||||
Db        647 AATCCACCCCGCTGCACTTCCCCGCAGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGC 706

Qy        457 TTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATG 516
              |||| |||||||||||| ||||| ||| | || ||||||||||||||||||||||||| ||
Db        707 TTCAAAATGGTGCAAATNTCCAANCACTTTATAATGGGGGCCTTATTCCTCTTCATANTG 766

Qy        517 CATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCA 576
              |||||||||||||||||||| ||     |||||||||||||||||||||||||||||||||
Db        767 CATGCTCTTTTGGTCATGCTAAAANTATCAATCTCCTTTTGCGACATGGTGCAGACCCCA 826

Qy        577 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 636
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 886

Qy        637 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 946

Qy        697 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 1006

Qy        757 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 1066

Qy        817 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 1126

Qy        877 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 1186

Qy        937 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 1246

Qy        997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1306

Qy       1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1366

Qy       1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1426

Qy       1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1486

Qy       1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1546

Qy       1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1606

Qy       1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1666

Qy       1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1667 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1726

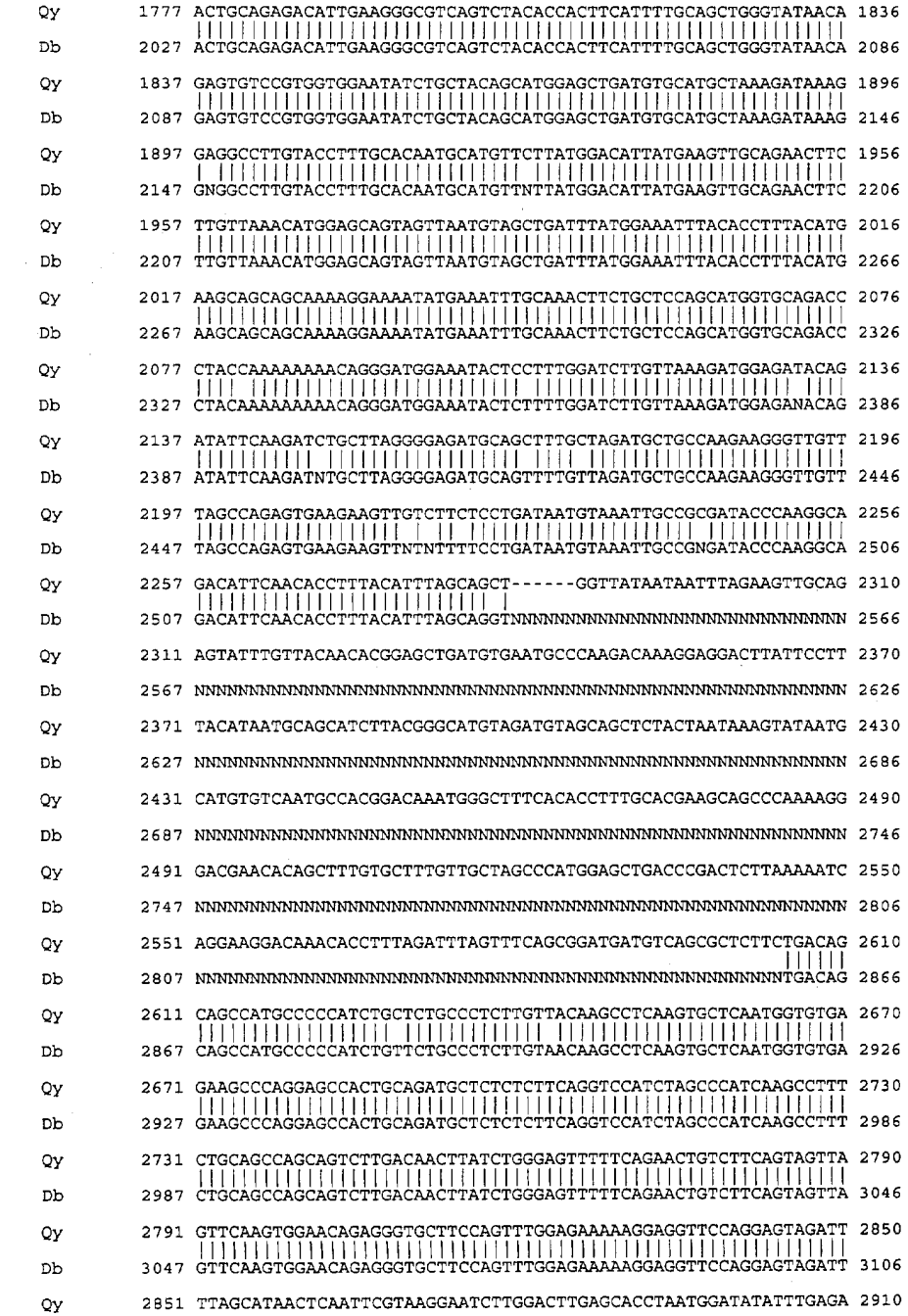
Qy       1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1727 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1786

Qy       1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1787 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1846

Qy       1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1847 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1906

Qy       1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1907 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1966

Qy       1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1967 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 2026
```



RESULT 18
AAC89706
ID AAC89706 standard; cDNA; 5005 BP.
XX
AC AAC89706;

XX
DT 08-MAR-2001 (first entry)
XX
DE Human adult T-cell leukaemia cDNA MO-BC-203.
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR P-PSDB; AAB48574.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 8; Page 82-84; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 5005 BP; 1492 A; 917 C; 1015 G; 1571 T; 0 U; 10 Other;

Query Match 72.7%; Score 2774.6; DB 4; Length 5005;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2810; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

```
Qy        997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 60

Qy       1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 120

Qy       1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 180

Qy       1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240

Qy       1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 300

Qy       1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 360

Qy       1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
              ||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db        361 AAAGAAAGCAAATATGTGAACTGTTGCTAAAAAAAGGATCAAACATCAATGAAAAGACTA 420

Qy       1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480

Qy       1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540

Qy       1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 600

Qy       1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 660

Qy       1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 720

Qy       1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 780

Qy       1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 840

Qy       1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 900

Qy       1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 960

Qy       1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2016
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1020

Qy       2017 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2076
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 1080

Qy       2077 CTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2136
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 1140

Qy       2137 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 1200

Qy       2197 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 1260

Qy       2257 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 2316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 1320

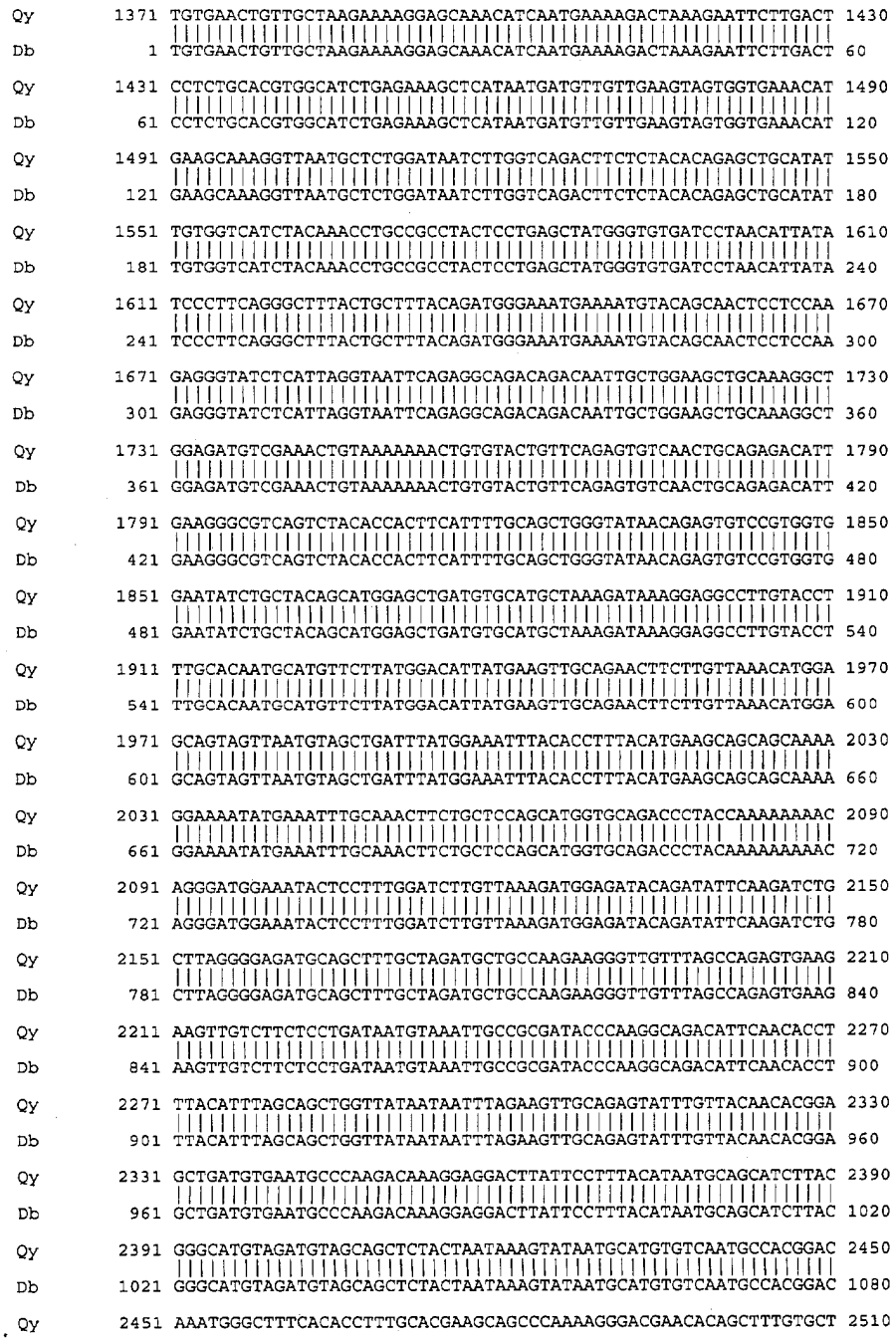
Qy       2317 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 2376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 1380

Qy       2377 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 2436
```

```
Db      1381 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 1440
Qy      2437 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 2496
Db      1441 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 1500
Qy      2497 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 2556
Db      1501 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 1560
Qy      2557 GACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCA 2616
Db      1561 GACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCA 1620
Qy      2617 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 2676
Db      1621 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 1680
Qy      2677 CAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 2736
Db      1681 CAGGAGCCACTGCAGATGCTCTCT-TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 1739
Qy      2737 CCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 2796
Db      1740 CCAGCAGTCTTGACAAC-TATCTGGGAG-TTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 1797
Qy      2797 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 2856
Db      1798 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 1857
Qy      2857 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 2916
Db      1858 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 1917
Qy      2917 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 2976
Db      1918 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 1977
Qy      2977 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 3036
Db      1978 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 2037
Qy      3037 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 3096
Db      2038 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 2097
Qy      3097 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 3156
Db      2098 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 2157
Qy      3157 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 3216
Db      2158 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 2217
Qy      3217 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 3276
Db      2218 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 2277
Qy      3277 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 3336
Db      2278 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 2337
Qy      3337 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3396
Db      2338 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 2397
Qy      3397 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 3456
Db      2398 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 2457
Qy      3457 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 3516
Db      2458 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 2517
Qy      3517 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 3576
Db      2518 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 2577
Qy      3577 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 3636
Db      2578 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 2637
Qy      3637 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 3696
Db      2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 2697
Qy      3697 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 3756
Db      2698 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 2757
Qy      3757 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db      2758 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2814
```

RESULT 19
AAF63919
ID AAF63919 standard; DNA; 2971 BP.
XX
AC AAF63919;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus SEQ ID NO: 92.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 152-153; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 2971 BP; 915 A; 555 C; 642 G; 859 T; 0 U; 0 Other;

Query Match 63.9%; Score 2439.8; DB 5; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;


Qy 1371 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1430
Db 1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60
Qy 1431 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1490
Db 61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120
Qy 1491 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
Db 121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180
Qy 1551 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1610
Db 181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240
Qy 1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1670
Db 241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300
Qy 1671 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1730
Db 301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360
Qy 1731 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1790
Db 361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420
Qy 1791 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1850
Db 421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480
Qy 1851 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1910
Db 481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540
Qy 1911 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1970
Db 541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600
Qy 1971 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2030
Db 601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660
Qy 2031 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2090
Db 661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720
Qy 2091 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2150
Db 721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780
Qy 2151 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2210
Db 781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840
Qy 2211 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2270
Db 841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900
Qy 2271 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2330
Db 901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960
Qy 2331 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2390
Db 961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020
Qy 2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2450
Db 1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080
Qy 2451 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2510
Db 1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140
Qy 2511 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2570
Db 1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200
Qy 2571 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630
Db 1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260
Qy 2631 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2690
Db 1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320
Qy 2691 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2750
Db 1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380
Qy 2751 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2810
Db 1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440
Qy 2811 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2870
Db 1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500
Qy 2871 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930
Db 1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560
Qy 2931 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2990
Db 1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620
Qy 2991 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3050
Db 1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680
Qy 3051 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3110
Db 1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740
Qy 3111 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3170
Db 1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800
Qy 3171 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3230
Db 1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860
Qy 3231 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3290
Db 1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920
Qy 3291 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3350
Db 1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980
Qy 3351 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3410
Db 1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040
Qy 3411 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3470
Db 2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100
Qy 3471 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3530
Db 2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160
Qy 3531 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3590

```
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3591 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3711 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy       3771 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443


RESULT 20
AAF63917
ID   AAF63917 standard; DNA; 3353 BP.
XX
AC   AAF63917;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone coding sequence SEQ ID NO: 88.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66285.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 1; Page 141-144; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 3353 BP; 1034 A; 616 C; 709 G; 994 T; 0 U; 0 Other;

  Query Match             63.9%;  Score 2439.8;  DB 5;  Length 3353;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2441;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1371 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1430
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1431 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1490
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1491 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1551 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1670
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1671 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1730
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1731 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1790
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1791 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1850
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1851 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1910
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1911 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1970
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1971 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2030
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2031 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2090
             |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720

Qy       2091 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2151 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2210
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2211 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2270
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2271 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2330
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2331 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2390
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2450
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2451 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2510
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140
```

```
Qy       2511 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2570
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy       2571 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2631 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2690
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy       2691 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy       2751 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy       2811 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy       2871 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy       2931 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2990
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2991 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy       3051 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy       3111 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy       3171 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860

Qy       3231 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy       3291 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy       3351 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy       3411 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy       3471 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3530
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy       3531 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3590
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3591 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3711 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

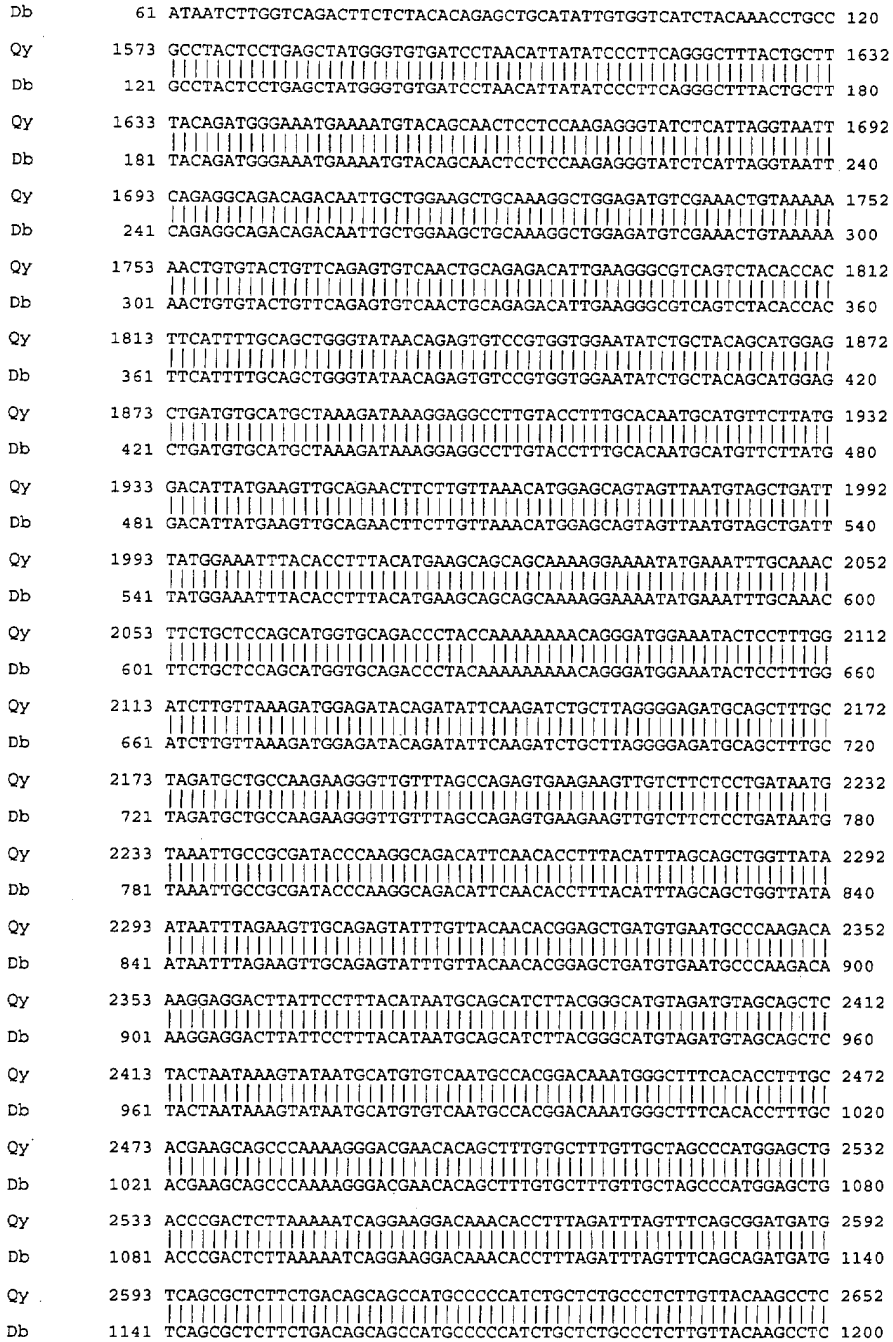
Qy       3771 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             |||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 21
AAF63918
ID AAF63918 standard; DNA; 3799 BP.
XX
AC AAF63918;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone coding sequence SEQ ID NO: 90.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66286.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 146-150; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 3799 BP; 1157 A; 694 C; 806 G; 1142 T; 0 U; 0 Other;

Query Match 61.8%; Score 2357.8; DB 5; Length 3799;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1453 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 1512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 60

Qy       1513 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 1572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 61 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 120
Qy 1573 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 1632
Db 121 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 180
Qy 1633 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 1692
Db 181 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 240
Qy 1693 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1752
Db 241 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 300
Qy 1753 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 1812
Db 301 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 360
Qy 1813 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 1872
Db 361 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 420
Qy 1873 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 1932
Db 421 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 480
Qy 1933 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 1992
Db 481 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 540
Qy 1993 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 2052
Db 541 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 600
Qy 2053 TTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGG 2112
Db 601 TTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGG 660
Qy 2113 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 2172
Db 661 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 720
Qy 2173 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 2232
Db 721 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 780
Qy 2233 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 2292
Db 781 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 840
Qy 2293 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2352
Db 841 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 900
Qy 2353 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2412
Db 901 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 960
Qy 2413 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2472
Db 961 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 1020
Qy 2473 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2532
Db 1021 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 1080
Qy 2533 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2592
Db 1081 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 1140
Qy 2593 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2652
Db 1141 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 1200
Qy 2653 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2712
Db 1201 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 1260
Qy 2713 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2772
Db 1261 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 1320
Qy 2773 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2832
Db 1321 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 1380
Qy 2833 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2892
Db 1381 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 1440
Qy 2893 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2952
Db 1441 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 1500
Qy 2953 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 3012
Db 1501 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 1560
Qy 3013 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 3072
Db 1561 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 1620
Qy 3073 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 3132
Db 1621 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 1680
Qy 3133 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 3192
Db 1681 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 1740
Qy 3193 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 3252
Db 1741 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 1800
Qy 3253 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 3312
Db 1801 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 1860
Qy 3313 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 3372
Db 1861 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 1920
Qy 3373 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 3432
Db 1921 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 1980
Qy 3433 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 3492
Db 1981 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 2040
Qy 3493 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 3552
Db 2041 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 2100
Qy 3553 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 3612
Db 2101 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 2160
Qy 3613 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 3672
Db 2161 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 2220
Qy 3673 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 3732
Db 2221 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 2280

```
Qy       3733 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 3792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 2340

Qy       3793 TACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||
Db       2341 TACTCCTTTGCTGAAAAAAAA 2361
```

RESULT 22
AAC85295
ID AAC85295 standard; cDNA; 3254 BP.
XX
AC AAC85295;
XX
DT 29-MAR-2001 (first entry)
XX
DE Mouse SPANK cDNA.
XX
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 308. .3254
FT /*tag= a
FT /product= "Murine SPANK"
FT /transl_except= (pos:1754,aa:Gly)
FT /note= "This codon has an apparent 2 nucleotide deletion,
FT which alters the reading frame"
FT /transl_except= (pos:2145,aa:DTQIGKLLIG)
FT /note= "The nucleotides encoding residues 614 to 623 are
FT absent, which alters the reading frame"
FT /transl_except= (pos:2602. .2606,aa:Met)
FT /note= "This codon has an apparent 2 nucleotide
FT insertion, which alters the reading frame"
XX
PN WO200077225-A1.
XX
PD 21-DEC-2000.
XX
PF 09-JUN-2000; 2000WO-US015926.
XX
PR 11-JUN-1999; 99US-0138957P.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Chi N, Lodish HF;
XX
DR WPI; 2001-091404/10.
DR P-PSDB; AAB47023.
XX
PT New insulin signaling protein SPANK, useful for reducing body mass,
PT glucose intolerance or insulin resistance and for preventing or treating
PT obesity-related and muscle-related diseases.
XX
PS Claim 3; Fig 5A; 65pp; English.
XX
CC This sequence represents the mouse SPANK cDNA. The SPANK protein
CC comprises 3 domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP
CC (poly adenosine diphosphate-ribose polymerase) catalytic domain; and (c)
CC an ANK domain composed of ankyrin repeats. SPANK is a cytosolic protein
CC which can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC perinuclear region of adipocytes. It is an effector in the insulin
CC signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC mass, reducing glucose intolerance or insulin resistance, for preventing
CC or treating obesity-related diseases or disorders, such as obesity,
CC cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC also useful for preventing or treating muscle-related diseases or
CC disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC Antibodies immunospecific for SPANK are useful for detecting the presence
CC of SPANK polypeptide in a biological sample
XX
SQ Sequence 3254 BP; 871 A; 746 C; 863 G; 774 T; 0 U; 0 Other;

Query Match 55.6%; Score 2122; DB 4; Length 3254;
Best Local Similarity 77.7%; Pred. No. 0;
Matches 2895; Conservative 0; Mismatches 275; Indels 558; Gaps 5;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||  |  |||| | ||||||||||||||  || ||||||  ||| ||||||||||
Db         82 CGCGCTGCCTCCACCGCGGTGGGGCAGCCGGGGCTCACGGAGCCGAGCCGGGGGCGCGCG 141

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
               | |||||  |||||||||  | ||||  || |||||||||  ||||||  ||||||||||
Db        142 GCGACGCGG-CCATGGGACCGTGCCGACTCTGGTGACAGCGGGGAGCGGAGCGGCCCGGG 200

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              |||||||||||||| || ||||||||||||||||||||||||||||||||| ||||||||||
Db        201 CCCTGAGCGCGTCTCCTGCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCCGGGCTCCTGC 260

Qy        181 TCCGGTTGCTGGCGCTGTTGCT---GGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCC 237
              | |  ||||||||||| ||||    ||| | |||||||||| |||||||||| || ||||
Db        261 TGCCTTTGCTGGCGCTGCTGCTGGCGGCGGCGGCGGCCGGATCATGTCTGGCCGCC 320

Qy        238 GCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCG 297
              |||||||||||||||| |||||||||||||||||| |||||||||||||||||||| |||
Db        321 GCTGCGCCGGCGGGGGCGCGGCCTGCGCGAGCGCCGGGGCCGAGGCCGTGGAGCCGTCCG 380

Qy        298 CCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGA 357
              |||| |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||
Db        381 CCCGCGAGCTGTTCGAGGCTTGCCGCAACGGGGACGTGGAGCGAGTCAAGAGGCTGGTGA 440

Qy        358 CGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCG 417
              |||| |||||||||||||||||||||||||||||||||||||| |||||||||| |||||||
Db        441 CGCCCGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAGTCCACCCCGCTCCACTTCG 500

Qy        418 CCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCC 477
              |||||||||||| ||||| || |||||||||||| | |||||||| |||||||||||| |
Db        501 CCGCAGGTTTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGC 560

Qy        478 AAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTG 537
              |||| |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        561 AAGCGCGTGATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTG 620

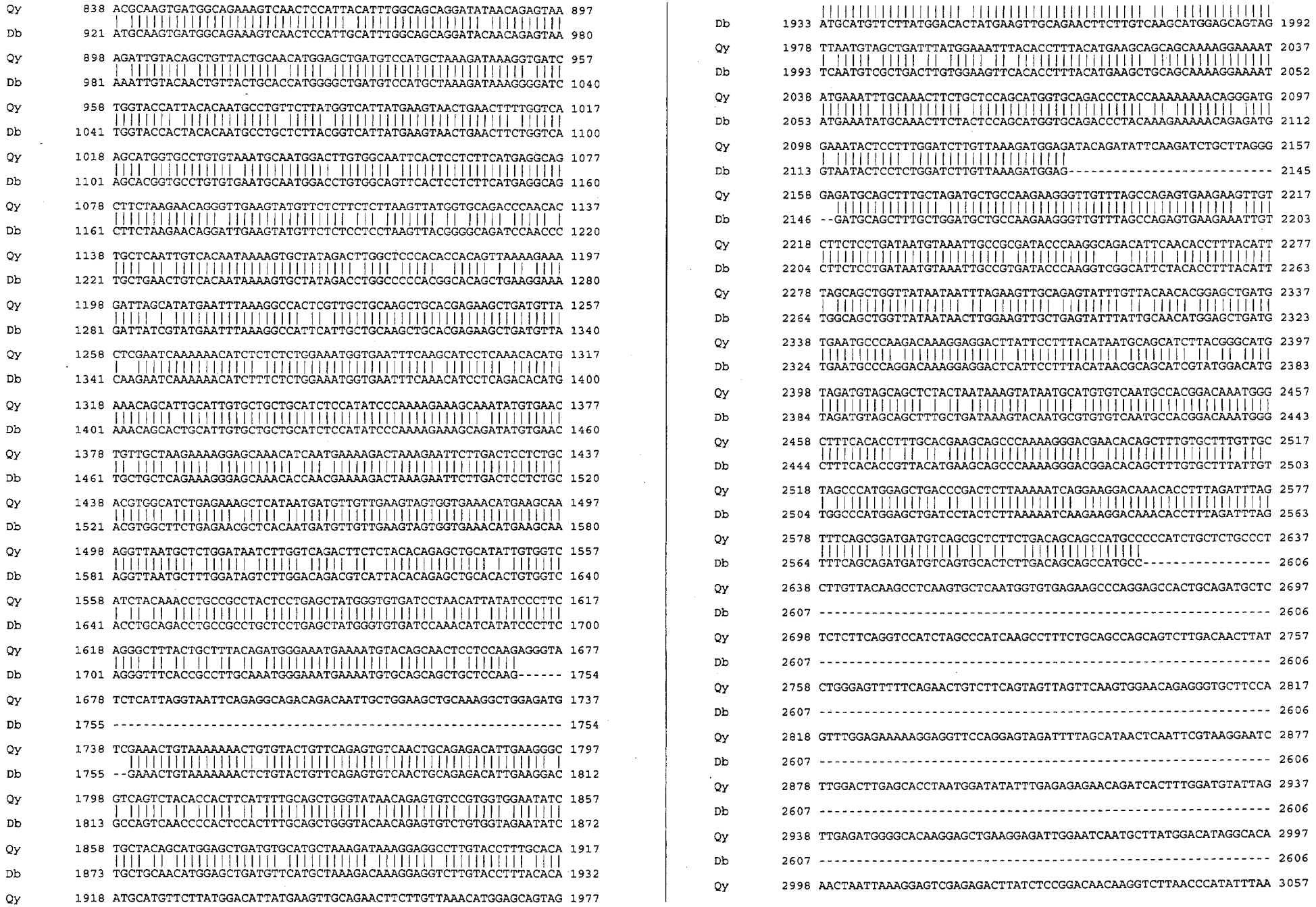
Qy        538 AAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATT 597
              |||||||||||| |||||||| ||| |||||||||||| |||||||| ||||||||||||
Db        621 AAGTAGTCAATCTTCTTTTGCAACACGGTGCAGACCCAAATGCTCGGGATAATTGGAATT 680

Qy        598 ATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTAC 657
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        681 ATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTAC 740

Qy        658 AGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAG 717
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 AGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAG 800

Qy        718 ATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTG 777
              | |||||||||||||| || || |||||||| |||||||||||||||||  | |||||||
Db        801 ACCCATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGAAAGATGAACTTCTGGAAAGTG 860

Qy        778 CCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCC 837
              |||||||||||||||| |||||||||||||| || || |||||||| ||||||||||||||
Db        861 CCAGGAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCC 920
```



```
Db       2607 ------------------------------------------------------------ 2606

Qy       3058 CTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGT 3117
                                  |||||||||| |||||||||||||||||||||||||||| |
Db       2607 --------------------GGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAAT 2646

Qy       3118 TTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATG 3177
              |||||||||||| |||||||||| ||||| || || |||||||||||||| ||||| |
Db       2647 TTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACG 2706

Qy       3178 CAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAAC 3237
              |||| ||  |||||||||||||||| |||||||||||||||||||||||||||||||||
Db       2707 CAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAAT 2766

Qy       3238 TATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCA 3297
              | |||||||||| || |||||||||||||||||||||||||||||||||||||||||| |
Db       2767 TGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCAA 2826

Qy       3298 ATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTG 3357
              | ||| | ||| |||||||||||||||||||||||||||| |||||||| || |||||||
Db       2827 ACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTG 2886

Qy       3358 ATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACT 3417
              |||||||||||||| |||||||||||||||| |||||||||||||| |||||||||||| ||||
Db       2887 ATGAAAGGCATGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACT 2946

Qy       3418 CTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACA 3477
              |||||||||||||||||||||| |||||||||||||| || || || |||||| ||||||
Db       2947 CTTCCAAAAGCAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTCCAATTCACA 3006

Qy       3478 AAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAA 3537
              |||||||||| |||||||||| |||||||||||||| ||||| || ||||||||||||||
Db       3007 AAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAA 3066

Qy       3538 AGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAG 3597
              |||||||| |||||||||| |||||||||||||||||||||||||| || |||||||| |
Db       3067 AGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGG 3126

Qy       3598 TCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAG 3657
              |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3127 TCACTGGCCGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAG 3186

Qy       3658 AACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCG 3717
              ||||||||||||||| ||| |||||||||||||| | |||||||||||||||||||| |
Db       3187 AACAGGCTTATCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAGGTATGGTTG 3246

Qy       3718 ATGGATAA 3725
              |||| | |
Db       3247 ATGGCTGA 3254
```

```
RESULT 23
AAF63838
ID   AAF63838 standard; DNA; 3984 BP.
XX
AC   AAF63838;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase1 coding sequence SEQ ID NO: 3.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66279.
XX
PT   New tankyrase2 polypeptides, useful by treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 1; Page 113-118; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 3984 BP; 1066 A; 1013 C; 1001 G; 904 T; 0 U; 0 Other;

  Query Match            49.4%;  Score 1885;  DB 5;  Length 3984;
  Best Local Similarity  72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;
```

```
Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | | || |  || |||     ||||   || || ||| |
Db        502 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 561

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              |||||  ||  |||||  |||||||||    || ||  ||||||||||||     | |  || ||
Db        562 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 621

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||        |||| ||| ||| |||| ||  | || ||||||||||| ||||||||||
Db        622 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 681

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              ||||  || || ||  ||| | || || ||||  ||||| |||||||| || |||||||||
Db        682 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 741

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              || || || || ||  |||||||||| || |||||||| |||||||| || || | |||
Db        742 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 801

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              | ||| | || || || || || ||||| | ||||| ||||| ||||| ||||| |||
Db        802 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 861

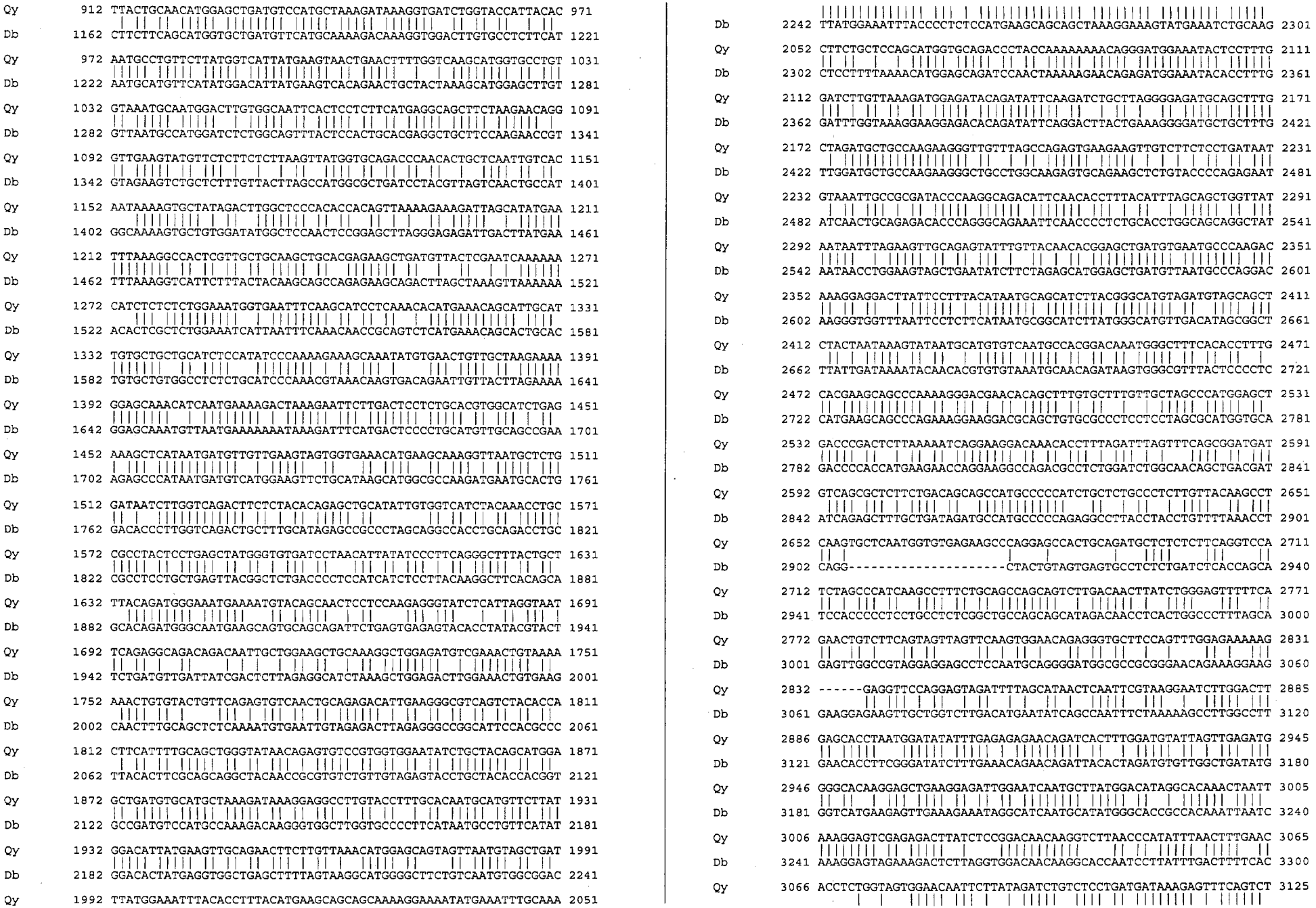
Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||| |||||||| ||||| ||||| |||||||||||| | ||||| |||||||||
Db        862 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 921

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              |||| ||| || || || ||||| |    |||| |||| | |||||||| || || |||
Db        922 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 981

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              || || ||||| |||||||| |||||||| ||||| |||||  ||| |||||||| |||
Db        982 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1041

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              |||||||||  ||||||  |||| || ||  |||||||  || ||||| ||||||||||
Db       1042 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1101

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              |||||||  |||||  ||||||  | |||||  || || ||||||||     || || |||||
Db       1102 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1161
```



```
Db       3301 TGTGTTAATCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCA 3360

Qy       3126 GTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGA 3185

Db       3361 GTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGC 3420

Qy       3186 ATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAA 3245

Db       3421 ATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAG 3480

Qy       3246 AGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGA 3305

Db       3481 CGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGC 3540

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365

Db       3541 ATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGA 3600

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425

Db       3601 CATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAA 3660

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485

Db       3661 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3720

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545

Db       3721 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3780

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605

Db       3781 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3840

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665

Db       3841 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3900

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708

Db       3901 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3943
```

```
RESULT 24
AAZ29627
ID   AAZ29627 standard; cDNA; 4134 BP.
XX
AC   AAZ29627;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human tankyrase cDNA.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   human testis library; recombinant tankyrase; telomere length homeostasis;
KW   skin atrophy; macular degeneration; atherosclerosis; gene therapy;
KW   ataxia telangiectasia; drug screening; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             6. .3989
FT                   /*tag=  a
FT                   /product= "Human tankyrase protein"
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   P-PSDB; AAY44402.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Claim 5; Page 111-112; 133pp; English.
XX
CC   The present cDNA sequence encodes human tankyrase, a TRF1 binding
CC   protein. This cDNA was isolated from a human testis library. Tankyrase is
CC   involved in regulation of telomere length through TRF1 and may directly
CC   modulate the effect of TRF1. This gene is used to produce recombinant
CC   tankyrase, or its fragments, and these are used to identify specific
CC   modulators which are potential drugs to control telomere length
CC   homeostasis for countering telomere shortening associated with aging
CC   (e.g. atrophy of skin, macular degeneration or atherosclerosis) or
CC   diseases such as ataxia telangiectasia. This can also be used to identify
CC   homologous genes and in gene therapy (including antisense or ribozyme
CC   expression). Transgenic animals having at least one disrupted endogenous
CC   allele for tankyrase are used in drug screening
XX
SQ   Sequence 4134 BP; 1112 A; 1041 C; 1034 G; 947 T; 0 U; 0 Other;
```

```
  Query Match           49.4%;  Score 1885;  DB 3;  Length 4134;
  Best Local Similarity  72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311

Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371

Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431

Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491

Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551

Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611

Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671

Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731

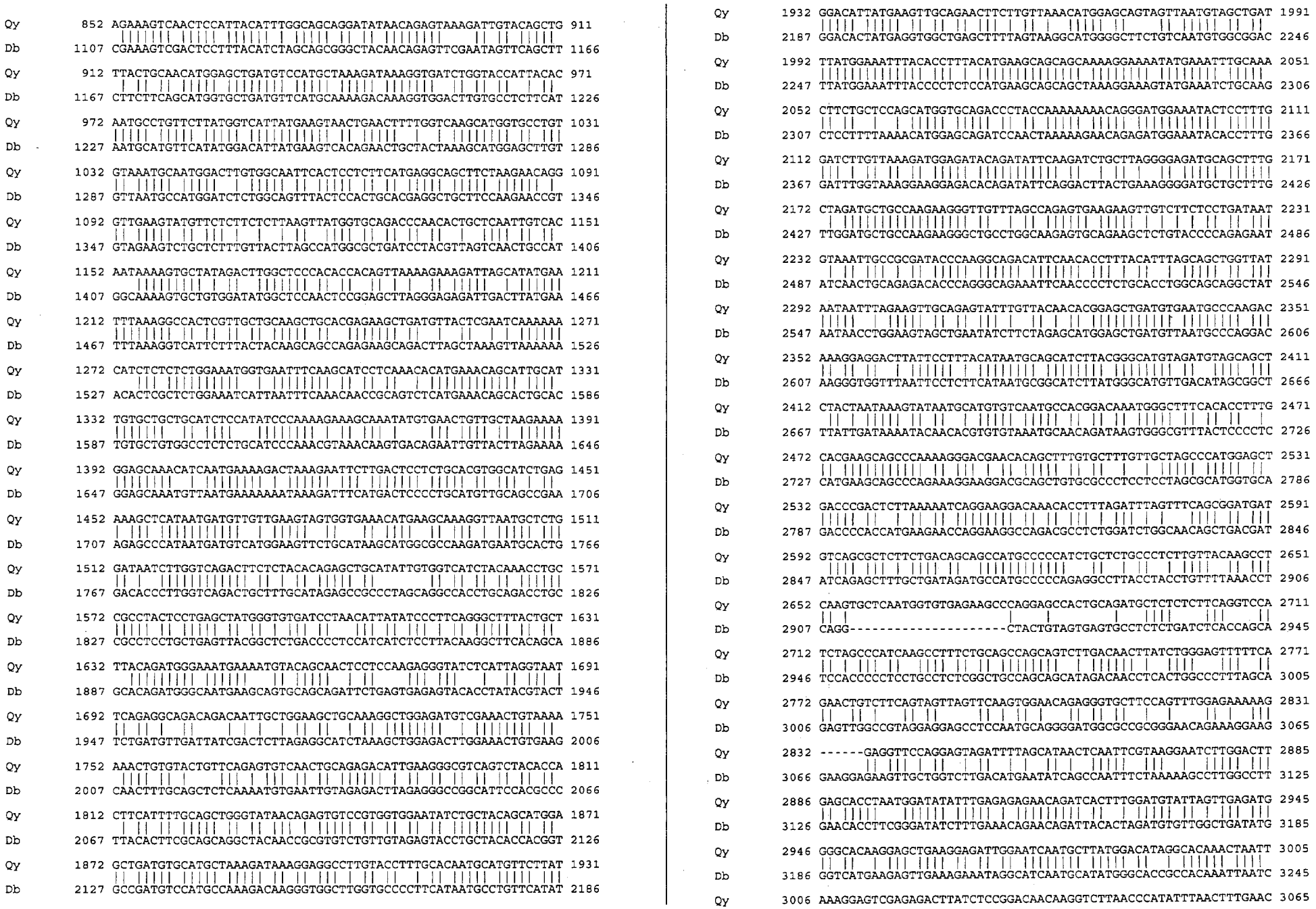
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791

Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851

Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106
```



```
                ||||||||| || |||||  |    ||||||||||    || || |||||  |||||  ||
Db       3246 AAAGGAGTAGAAAGACTCTTAGGTGGACAACAAGGCACCAATCCTTATTTGACTTTTCAC 3305

Qy       3066 ACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCT 3125
                |  |||||  |||  |  | |||||  ||||  ||  ||||||||  | ||||||
Db       3306 TGTGTTAATCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCA 3365

Qy       3126 GTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGA 3185
                |||||  |||||||||||||||||  |||||||  ||||||||||  |||  ||||  || ||
Db       3366 GTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGC 3425

Qy       3186 ATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAA 3245
                ||||||||||||||||||  |    |||||  ||  |||    ||||||||   |  ||||
Db       3426 ATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAG 3485

Qy       3246 AGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGA 3305
                | | |   |||||    || |||||  |||||  ||  ||  |||||||||   ||||||  ||
Db       3486 CGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGC 3545

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365
                |||  | ||||||||  ||||||||   | |||||  |||||  ||  |||||  |||||||||   |
Db       3546 ATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGA 3605

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425
                |||||  ||||||||  ||  |||||||||  ||  ||  ||||||||||||||||||  ||  |||
Db       3606 CATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAA 3665

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485
                |||||  ||||||||  ||||||||||||||||||  ||  ||  ||  ||    |||||  |||||
Db       3666 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3725

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545
                ||  ||  ||  ||  ||||  ||   ||||||  ||   |  ||  |||  ||  ||  ||
Db       3726 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3785

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
                ||||||||  ||  |  ||||||||||  ||  |  ||||||||  ||  ||||||||||  ||||
Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
                ||  ||  ||  ||  |||||  ||  ||||  ||||||||||  ||  ||||||||||||||||
Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
                ||  ||  ||||||  |  ||  ||||||||||  ||||  |||  ||||
Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948
```

RESULT 25
AAC66826
ID AAC66826 standard; DNA; 4134 BP.
XX
AC AAC66826;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase I coding sequence SEQ ID NO: 7.
XX
KW Human; tankyrase II; telomere length; signal transduction; ds.
XX
OS Homo sapiens.
XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009558.
XX
PR 09-APR-1999; 99US-0128577P.
PR 13-APR-1999; 99US-0129123P.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 1; Fig 6; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase I coding
CC sequence. The invention relates to the isolation of the protein and
CC coding sequences of human tankyrase II. This protein is thought to be
CC involved in signal transduction in the cell, and to have binding activity
CC for other telomere-associated proteins. It is possible that it plays a
CC role in the regulation of telomere length, thus affecting the replicative
CC ability of the cell. The protein is useful for ribosylating target
CC proteins, for determining tankyrase II binding activity in a sample, and
CC for modulating telomere length in a cell
XX
SQ Sequence 4134 BP; 1112 A; 1041 C; 1034 G; 947 T; 0 U; 0 Other;

Query Match 49.4%; Score 1885; DB 3; Length 4134;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps 2;

```
Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
                || ||  |  |  |  | | ||  |  || |||     ||||   || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
                |||||  || |||||  |||||||||    || ||  ||||||||||||     | |  || ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
                ||          ||||  ||| ||| |||| ||  | || ||||||||||| ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
                 |||| || || || ||| | || || ||||  |||||  |||||||| || |||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
                || || || || || |||||||||| || |||||||| |||||||| || || | |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
                | ||| | || || || || |||||  | |||||  ||||| |||||  ||||| |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
                |||||||| |||||||| ||||| ||||| ||||||||||| | ||||| ||||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

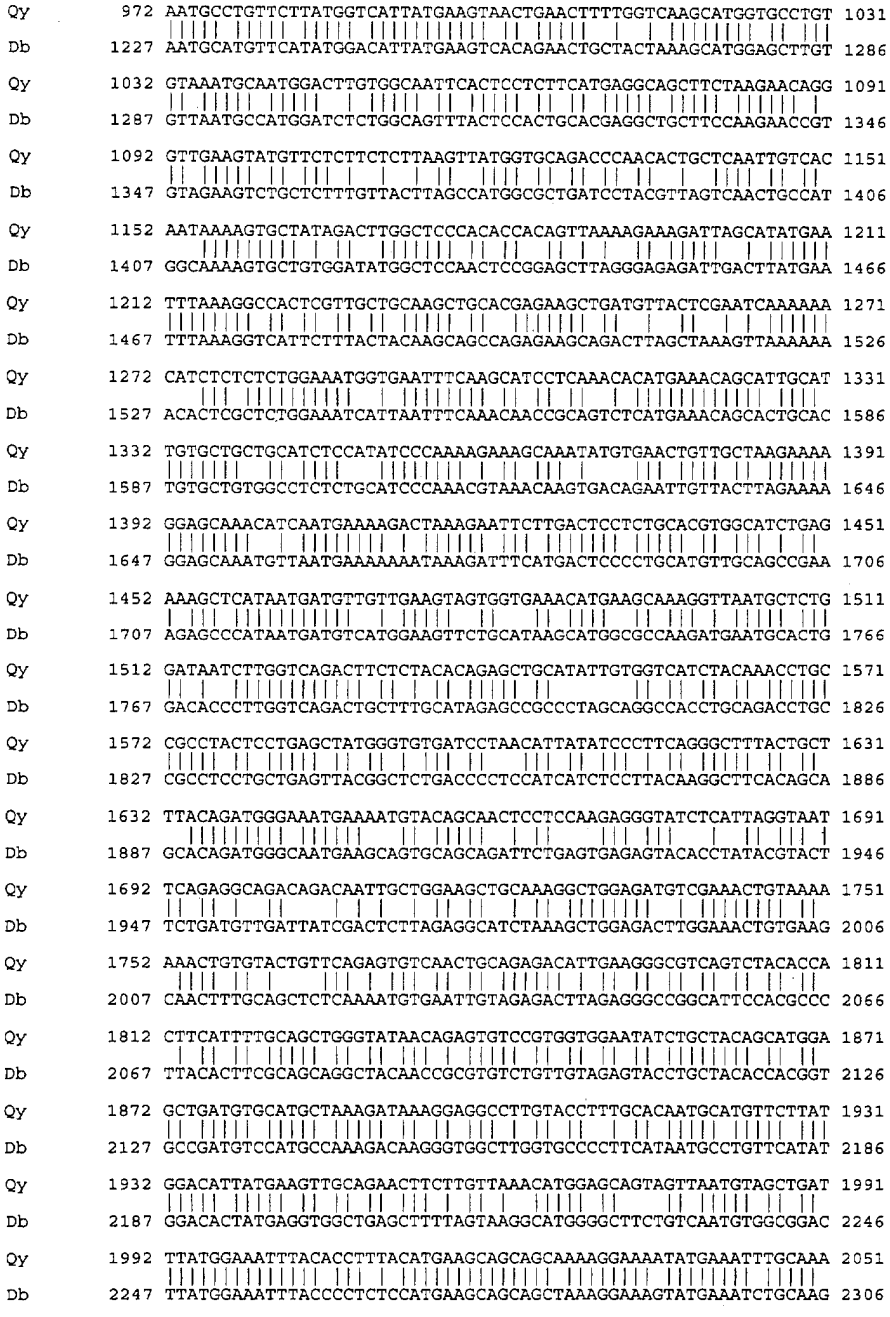
Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
                |||| ||| || || || |||||  |||  ||||  |||||  |||||||  || || |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

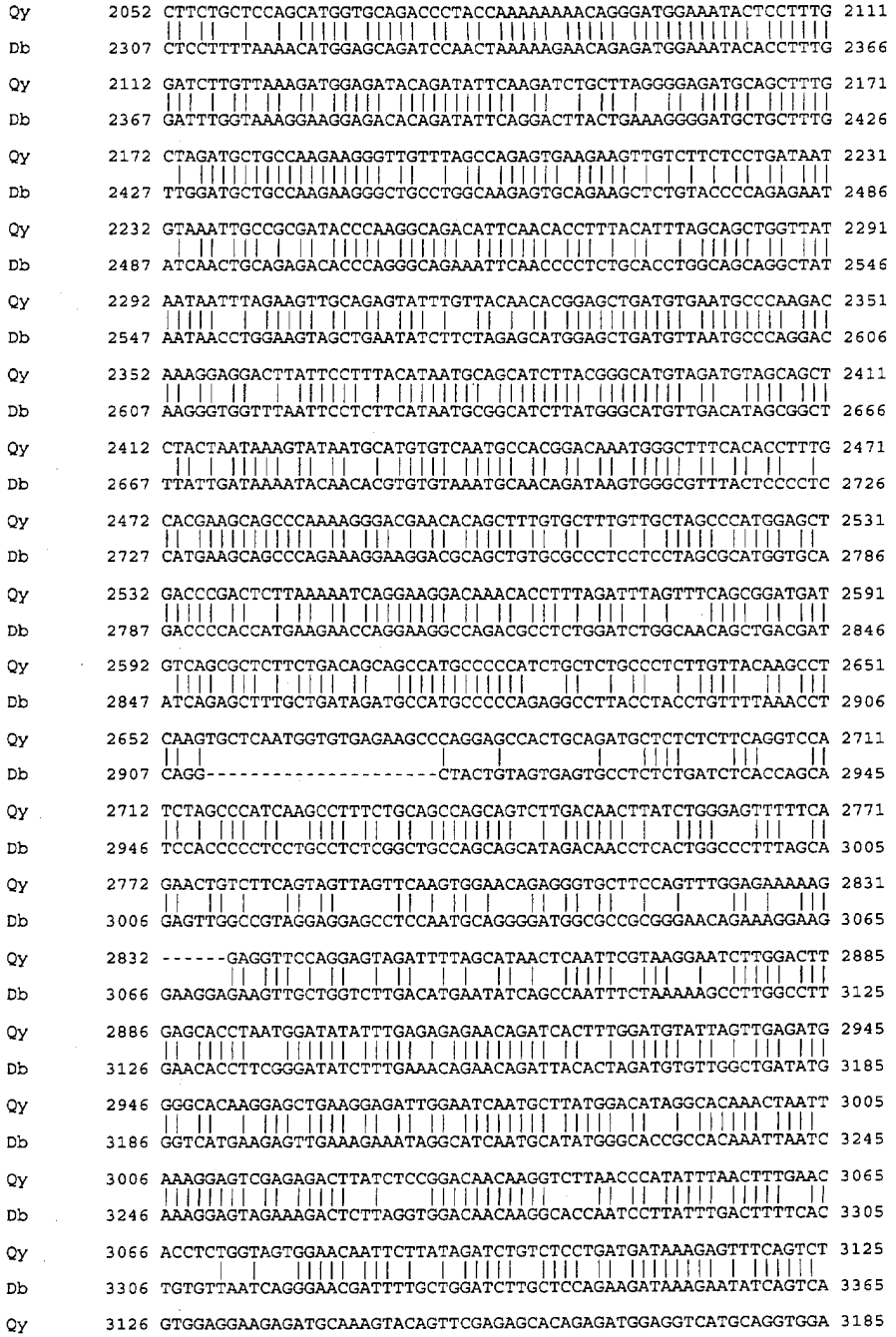
Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
                || || |||||  |||||||| |||||||  ||||||  |||||  |||  |||||||| |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
                |||||||| | ||||||  ||||  ||  ||  ||||||  ||  |||||  ||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
                |||||||  |||||  ||||||  | |||||  ||  ||  ||||||||    ||  ||  |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
                | ||  ||  |||||  ||||||||  |||||  |||||  |||||||    ||  ||  ||   | ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226
```





```
Db       3366 GTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGC 3425

Qy       3186 ATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAA 3245

Db       3426 ATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAG 3485

Qy       3246 AGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGA 3305

Db       3486 CGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGC 3545

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365

Db       3546 ATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGA 3605

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425

Db       3606 CATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAA 3665

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485

Db       3666 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3725

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545

Db       3726 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3785

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605

Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665

Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708

Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948
```

```
RESULT 26
ABK48629
ID   ABK48629 standard; cDNA; 4134 BP.
XX
AC   ABK48629;
XX
DT   12-AUG-2002  (first entry)
XX
DE   Human tankyrase cDNA.
XX
KW   Human; tankyrase; TRF1 interacting ankyrin; ss; gene; cytostatic;
KW   ADP ribosylation; telomere; nuclear localisation signal; NLS;
KW   telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW   drug development; aging; ataxia telangiectasia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             6. .3989
FT                   /*tag=  a
FT                   /product= "Tankyrase"
FT   3'UTR           3990. .4134
FT                   /*tag=  b
XX
PN   WO200229017-A2.
XX
PD   11-APR-2002.
XX
PF   25-SEP-2001; 2001WO-US029931.
XX
PR   03-OCT-2000; 2000US-00678411.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2002-463230/49.
DR   P-PSDB; AAU79537.
XX
PT   Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT   to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT   ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT   -ribosylation activity.
XX
PS   Disclosure; Page 127-129; 151pp; English.
XX
CC   The invention discloses a nuclear localisation signal (NLS)-telomeric
CC   protein tankyrase that comprises a NLS and at least a portion of a
CC   tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC   factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC   Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC   TRF1, which is a negative regulator of telomere length maintenance, and
CC   in so doing, counteracts its effect and play a role in lifecycle of a
CC   cell. The protein is useful for identifying agents that modulate the
CC   binding of TRF1 to its DNA binding site, for identifying agents that
CC   modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC   identifying the agents that modulate telomere length and for identifying
CC   agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC   the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC   would hinder telomere lengthening and so is useful in the treatment of
CC   cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC   telomere lengthening and so could be useful for extending the lifetime of
CC   the cell, especially self-renewing cells such as bone marrow cells,
CC   peripheral blood, T and B cells. Tankyrases are also useful for
CC   developing drugs that can counteract the telomere shortening associated
CC   with aging and certain diseases such as ataxia telangiectasia. The
CC   sequence presented is the human tankyrase cDNA
XX
SQ   Sequence 4134 BP; 1112 A; 1041 C; 1034 G; 947 T; 0 U; 0 Other;
```

```
  Query Match           49.4%;  Score 1885;  DB 6;  Length 4134;
  Best Local Similarity  72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311

Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371

Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431

Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491

Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551

Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611

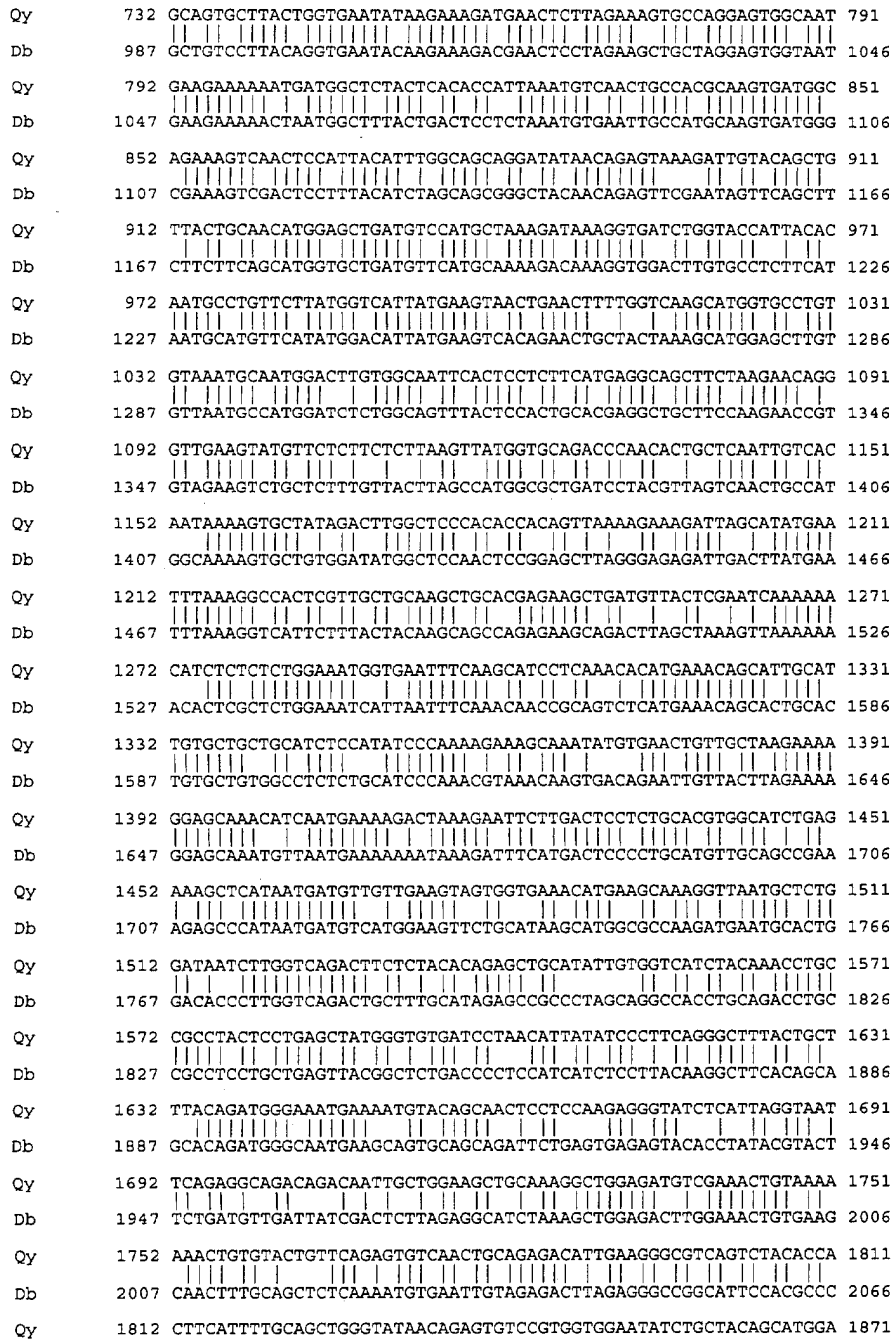
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

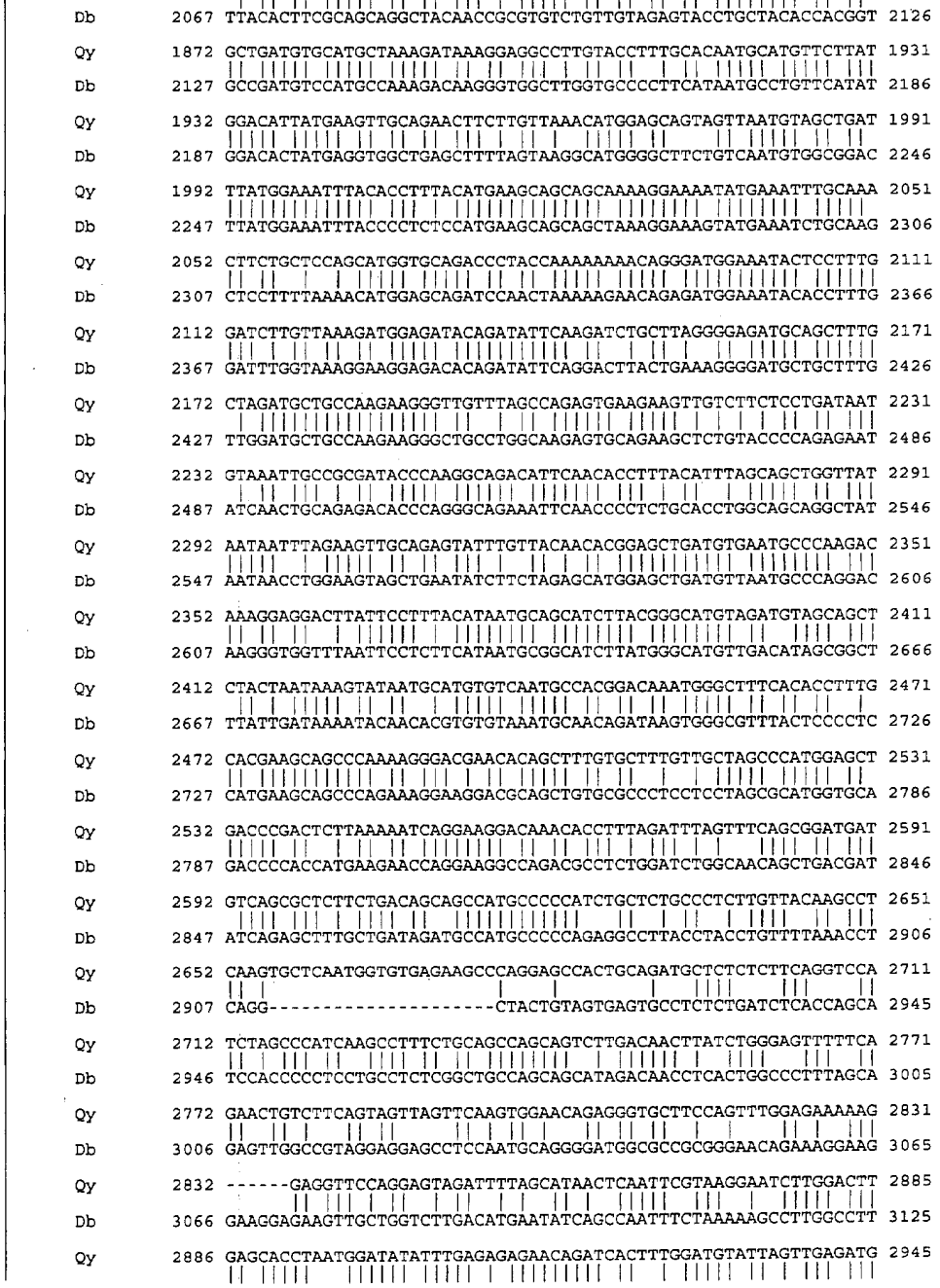
Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671

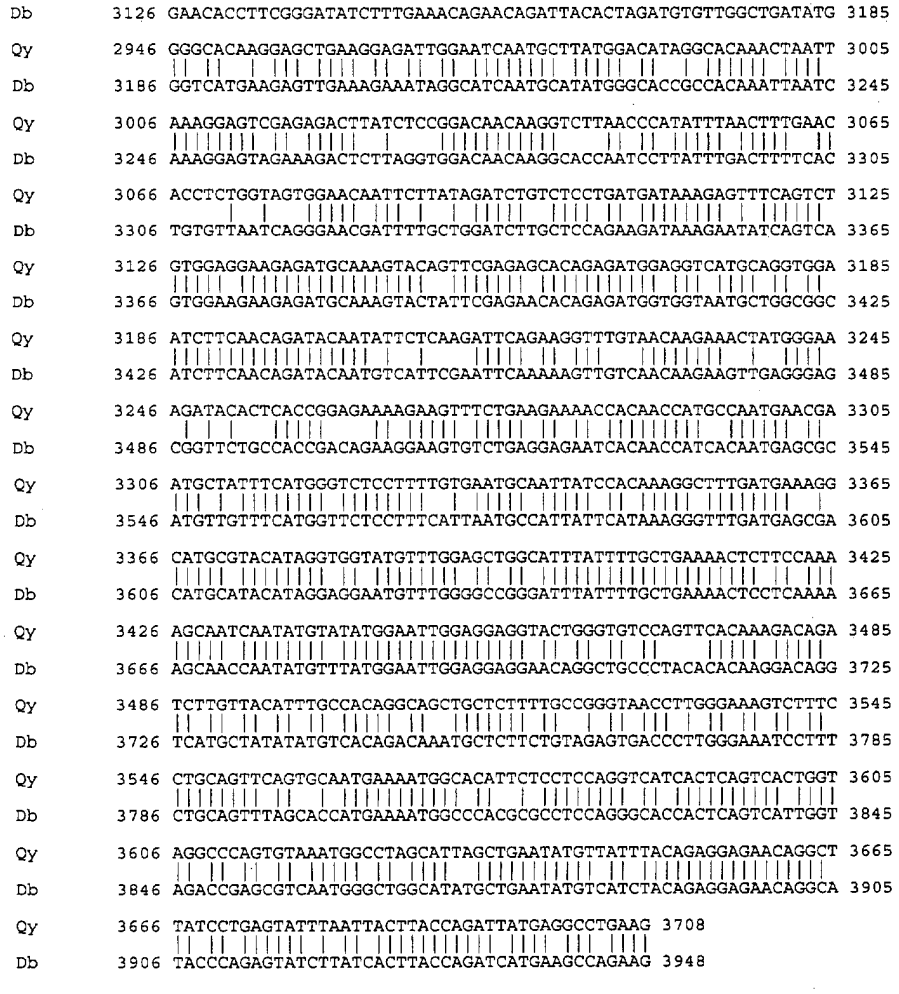
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731

Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986
```

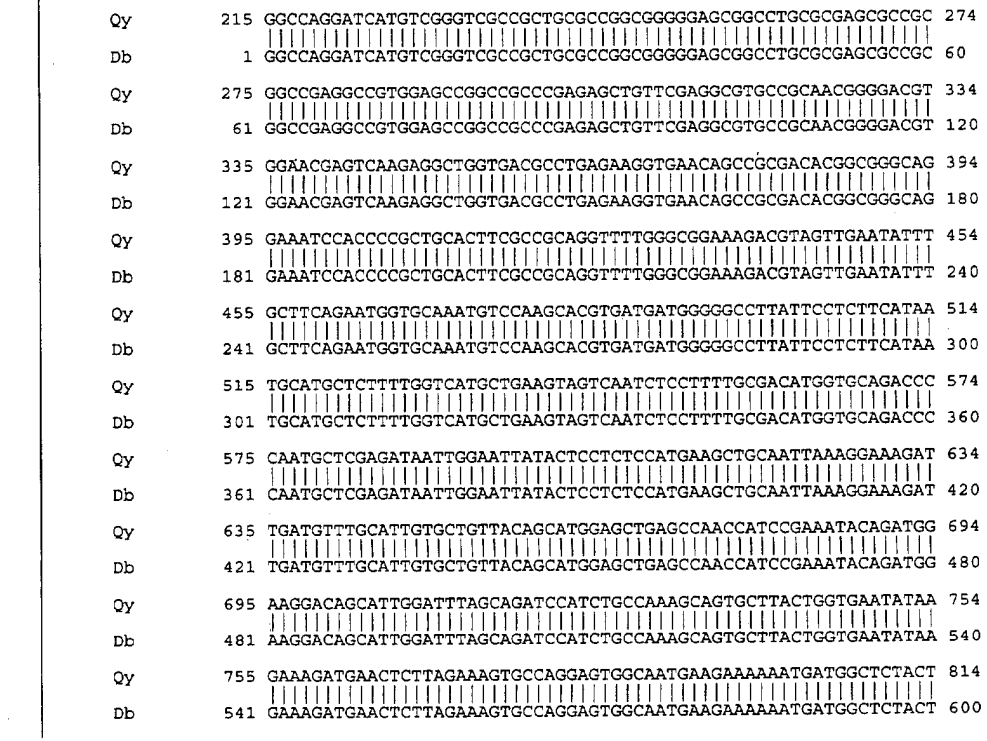






RESULT 27
AAF63925
ID AAF63925 standard; DNA; 1568 BP.
XX
AC AAF63925;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone coding sequence SEQ ID NO: 98.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66287.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 154-156; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1568 BP; 460 A; 337 C; 387 G; 384 T; 0 U; 0 Other;

Query Match 41.1%; Score 1568; DB 5; Length 1568;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840

Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900

Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960

Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020

Qy       1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080

Qy       1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140

Qy       1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200

Qy       1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260

Qy       1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320

Qy       1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380

Qy       1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440

Qy       1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500

Qy       1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560

Qy       1775 CAACTGCA 1782
              ||||||||
Db        1561 CAACTGCA 1568
```

```
RESULT 28
ADE53995
ID   ADE53995 standard; cDNA; 4435 BP.
XX
AC   ADE53995;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human prostate cancer cDNA #342.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.
XX
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.
PA   (PEAR/) PEARSON C I.
XX
PI   Faris M,  Pearson CI;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 5; SEQ ID NO 342; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
CC   cancer cDNA of the invention.
XX
SQ   Sequence 4435 BP; 1352 A; 757 C; 851 G; 1473 T; 0 U; 2 Other;

  Query Match            40.2%;  Score 1533.8;  DB 9;  Length 4435;
  Best Local Similarity  97.1%;  Pred. No. 0;
  Matches 1562;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
              || || |||| |  | |  ||| |  | |        | |  |||| |  | | |
Db        289 GTTAACAAGTGGAAGTATTATGACATAGGACTAAAAAAGAAAAAACAAGTAAAGTAATTT 348

Qy       2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
                ||   |  |||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TACTTCTTTTCTAAATTAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 408

Qy       2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 468

Qy       2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 528

Qy       2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 588
```

```
Qy       2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 648

Qy       2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 708

Qy       2625 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 768

Qy       2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 828

Qy       2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 888

Qy       2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 948

Qy       2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 1008

Qy       2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 1068

Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 1128

Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 1188

Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 1248

Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1249 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 1308

Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 1368

Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 1428

Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1429 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 1488

Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1489 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 1548

Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1549 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 1608

Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1609 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 1668

Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 1728

Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1729 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 1788

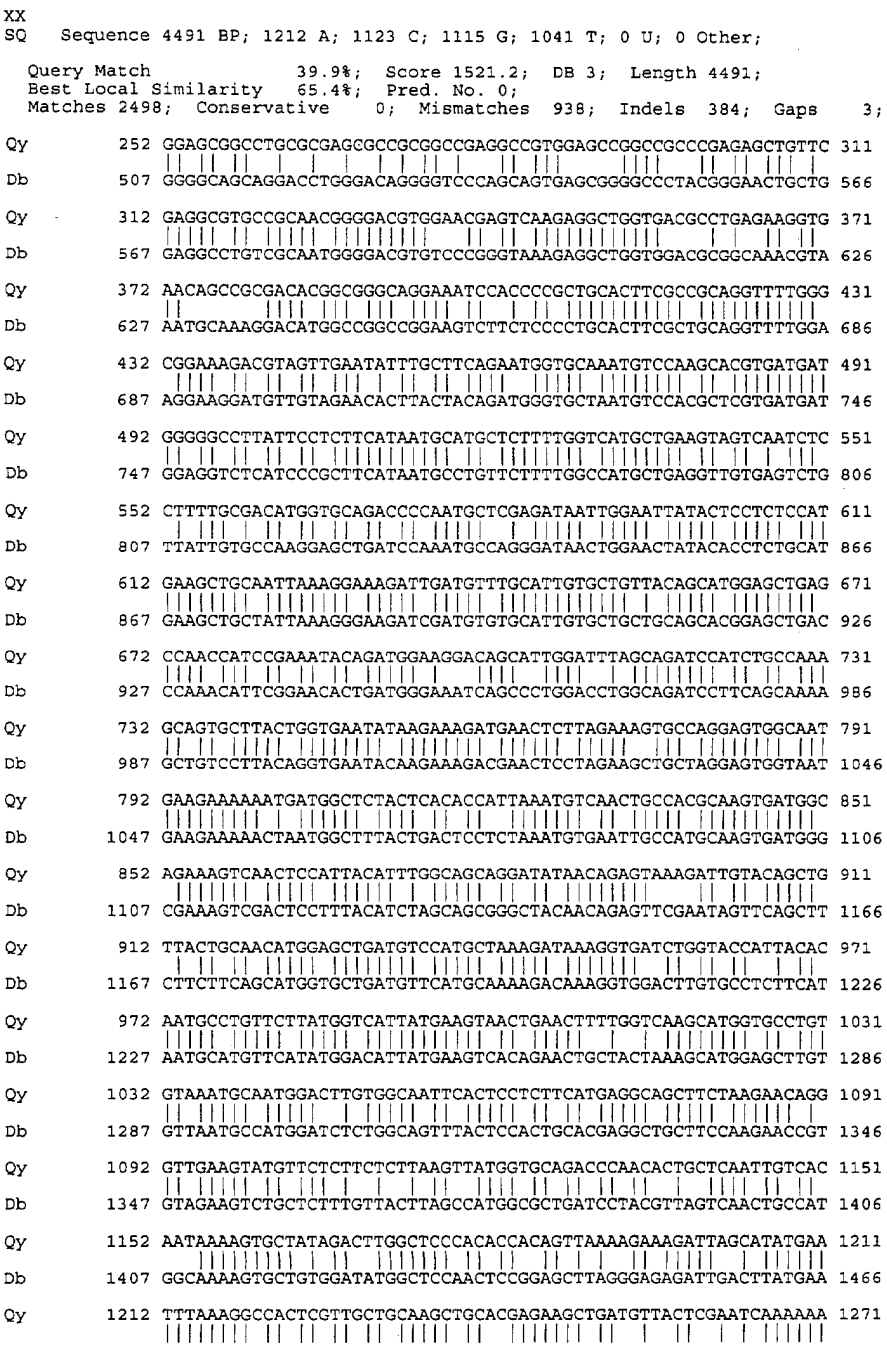
Qy       3705 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3764
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 1848

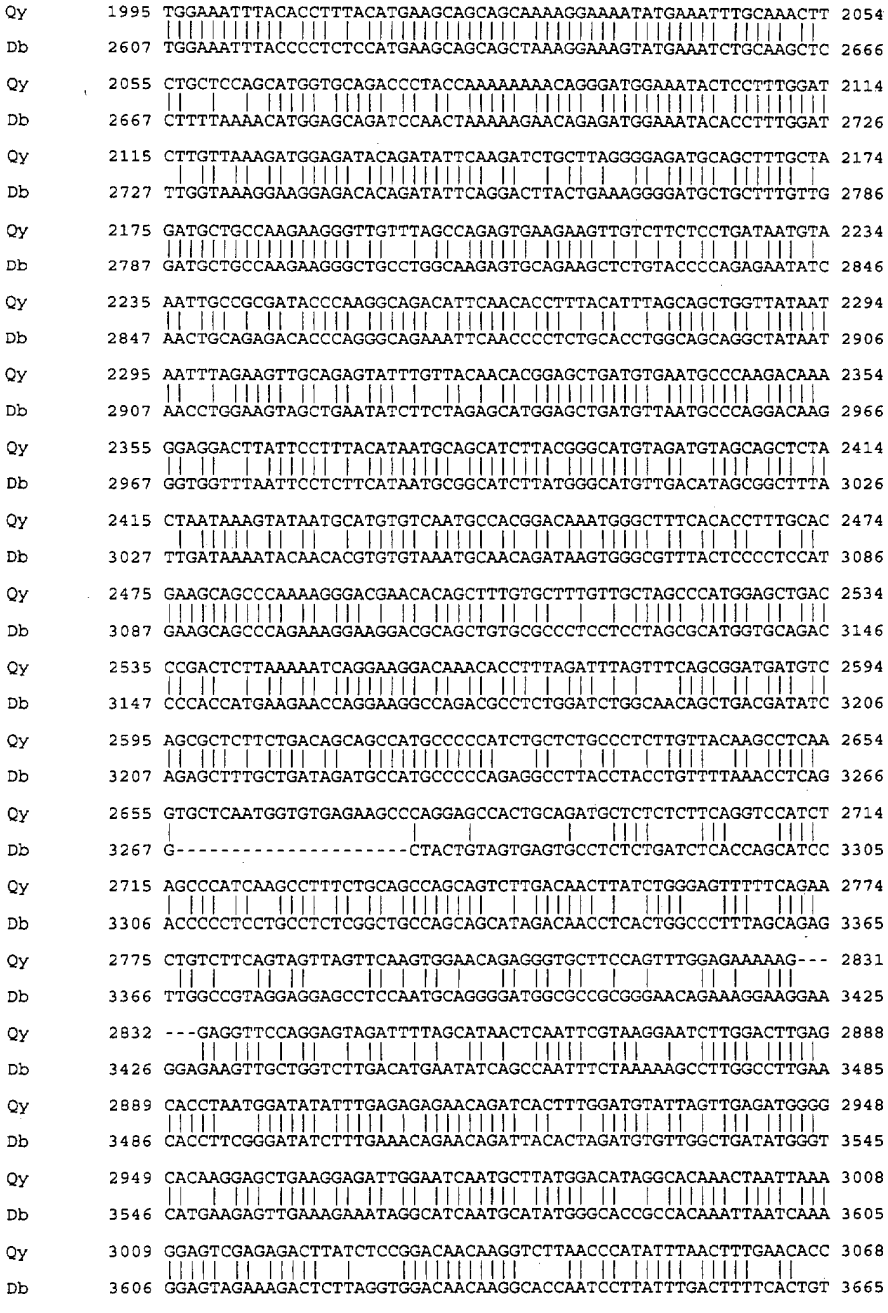
Qy       3765 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1849 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 1897
```

RESULT 29
AAZ29632
ID AAZ29632 standard; DNA; 4491 BP.
XX
AC AAZ29632;
XX
DT 22-MAR-2000 (first entry)
XX
DE Human truncated tankyrase-1 DNA.
XX
KW Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW recombinant tankyrase; telomere length homeostasis; skin atrophy;
KW macular degeneration; atherosclerosis; gene therapy;
KW ataxia telangiectasia; drug screening; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 6. .2027
FT /*tag= a
FT /product= "Human truncated tankyrase"
XX
PN WO9964606-A1.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-US012968.
XX
PR 10-JUN-1998; 98US-00095225.
PR 17-AUG-1998; 98US-00135233.
PR 19-NOV-1998; 98US-00196387.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2000-116549/10.
DR P-PSDB; AAY44403.
XX
PT New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT length, used to identify modulators, e.g. for inhibiting growth of
PT cancer.
XX
PS Disclosure; Page 119-120; 133pp; English.
XX
CC The present sequence encodes a truncated human tankyrase, a TRF1 binding
CC protein. This DNA comprises the coding sequence for amino acid residues 1
CC -640 of human tankyrase. Tankyrase is involved in regulation of telomere
CC length through TRF1 and may directly modulate the effect of TRF1. This
CC gene is used to produce recombinant tankyrase, or its fragments, and
CC these are used to identify specific modulators which are potential drugs
CC to control telomere length homeostasis for countering telomere shortening
CC associated with aging (e.g. atrophy of skin, macular degeneration or
CC atherosclerosis) or diseases such as ataxia telangiectasia. This can also
CC be used to identify homologous genes and in gene therapy (including
CC antisense or ribozyme expression). Transgenic animals having at least one
CC disrupted endogenous allele for tankyrase are used in drug screening

XX
SQ Sequence 4491 BP; 1212 A; 1123 C; 1115 G; 1041 T; 0 U; 0 Other;

Query Match 39.9%; Score 1521.2; DB 3; Length 4491;
Best Local Similarity 65.4%; Pred. No. 0;
Matches 2498; Conservative 0; Mismatches 938; Indels 384; Gaps 3;







```
RESULT 30
ABK48634
ID   ABK48634 standard; cDNA; 4491 BP.
XX
AC   ABK48634;
XX
DT   12-AUG-2002  (first entry)
XX
DE   Truncated human tankyrase #1 cDNA.
XX
KW   Human; tankyrase; TRF1 interacting ankyrin; ss; gene; cytostatic;
KW   ADP ribosylation; telomere; nuclear localisation signal; NLS;
KW   telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW   drug development; aging; ataxia telangiectasia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             6. .2027
FT                   /*tag=  a
FT                   /product= "Truncated tankyrase #1"
XX
PN   WO200229017-A2.
XX
PD   11-APR-2002.
XX
PF   25-SEP-2001; 2001WO-US029931.
XX
```

PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR P-PSDB; AAU79538.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 134-138; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC sequence presented is the truncated human tankyrase #1 cDNA. Note:
CC Nucleotides 1-1925 are the same as the wild-type human tankyrase sequence
CC (ABK48629)
XX
SQ Sequence 4491 BP; 1212 A; 1123 C; 1115 G; 1041 T; 0 U; 0 Other;

Query Match 39.9%; Score 1521.2; DB 6; Length 4491;
Best Local Similarity 65.4%; Pred. No. 0;
Matches 2498; Conservative 0; Mismatches 938; Indels 384; Gaps 3;

```
Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

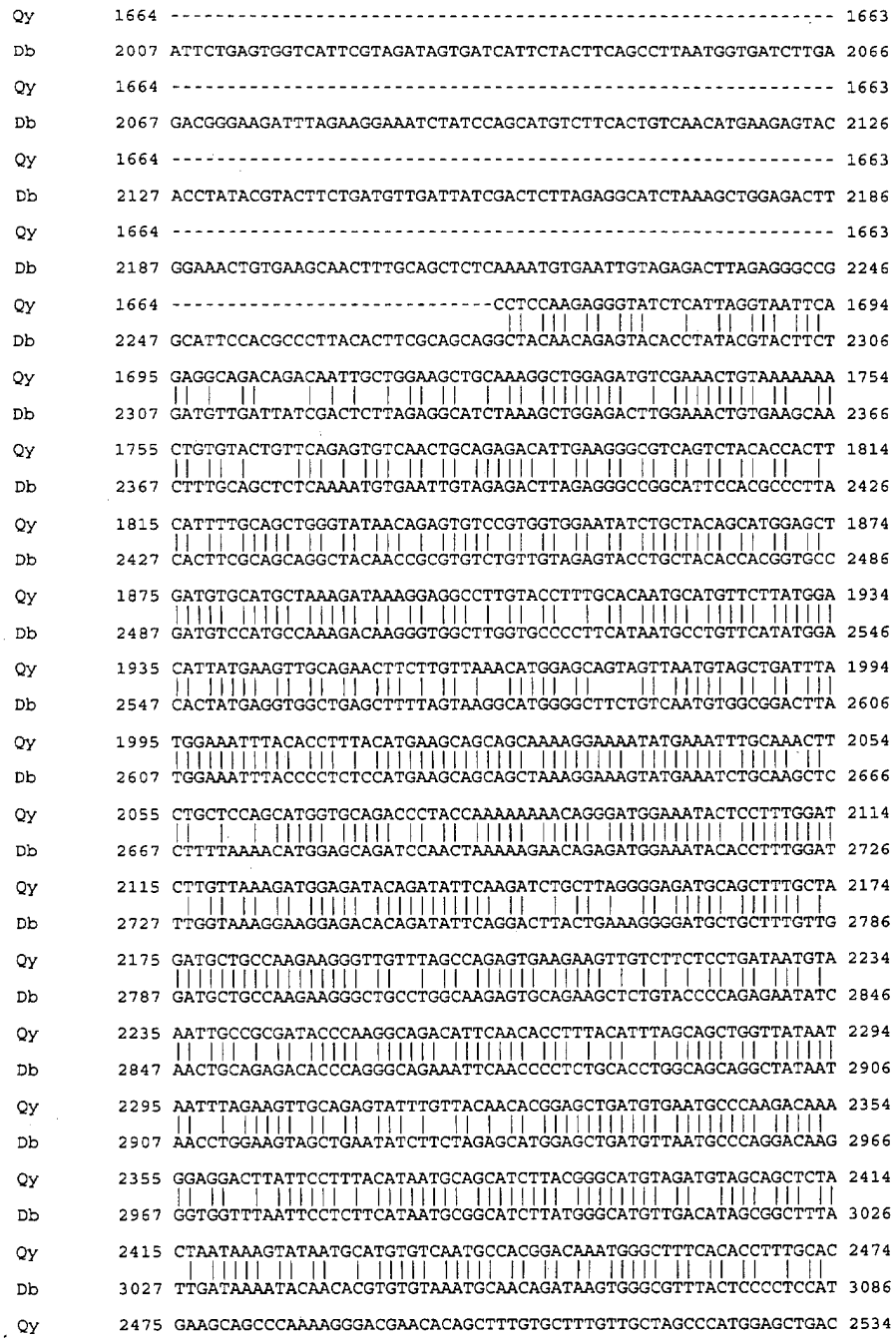
Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

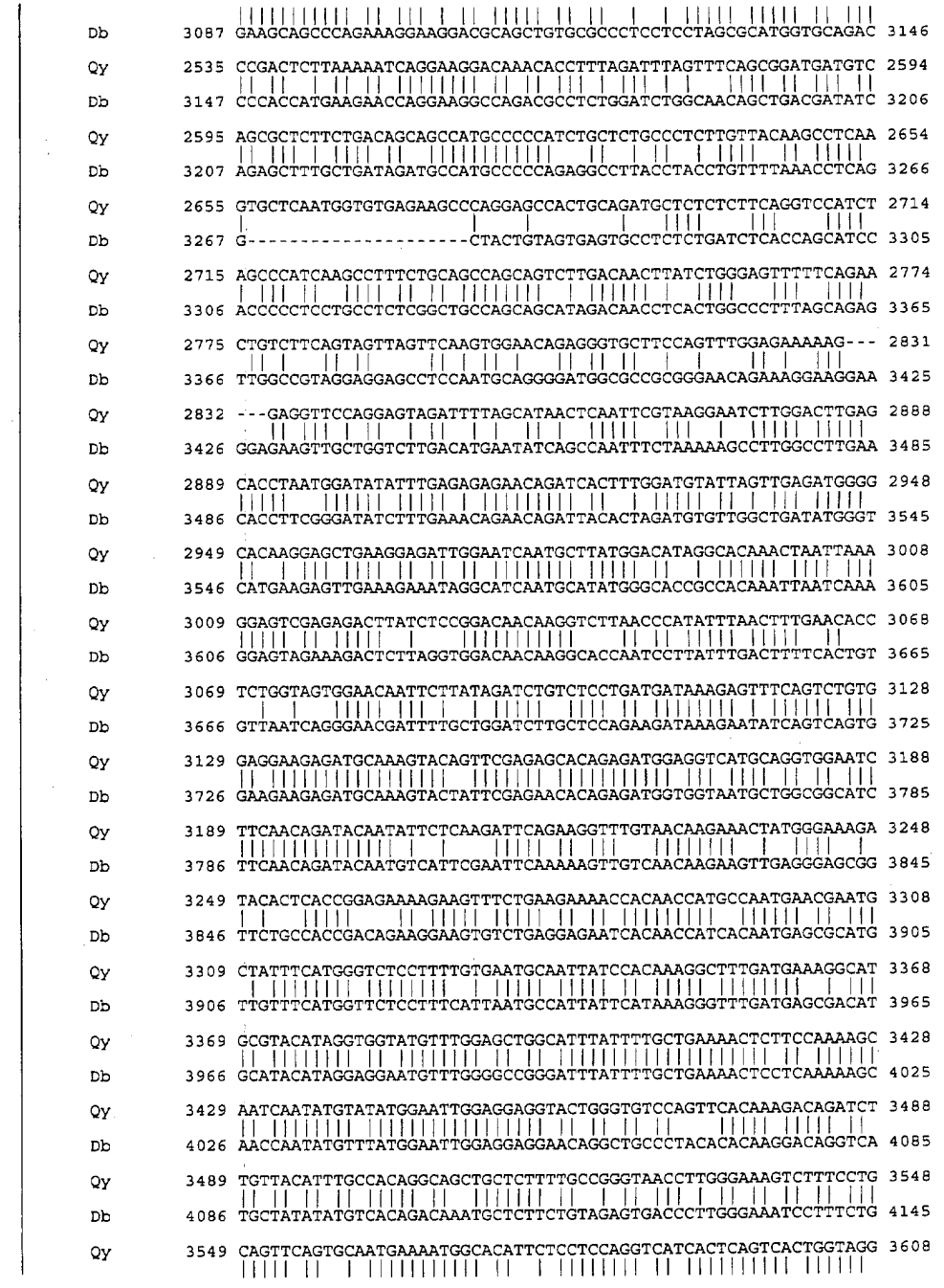
Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db       1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db       1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACT---------------------------- 1663
Db       1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGTGAGTTACGGCTCTGACCCC 1946

Qy       1664 ------------------------------------------------------------ 1663
Db       1947 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 2006
```





```
Db       4146 CAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGA 4205

Qy       3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668
             || || || ||||| || |||| |||||||||||| || ||||||||||||||||| ||
Db       4206 CCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATAC 4265

Qy       3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
              || |||||| | || ||||||||||| |||| ||| ||||
Db       4266 CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4305


RESULT 31
AAZ29633
ID   AAZ29633 standard; DNA; 4657 BP.
XX
AC   AAZ29633;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human truncated tankyrase-2 DNA.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   recombinant tankyrase; telomere length homeostasis; skin atrophy;
KW   macular degeneration; atherosclerosis; gene therapy;
KW   ataxia telangiectasia; drug screening; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             6. .2855
FT                   /*tag=  a
FT                   /product= "Human truncated tankyrase"
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   P-PSDB; AAY44404.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Disclosure; Page 123-125; 133pp; English.
XX
CC   The present sequence encodes a truncated human tankyrase, a TRF1 binding
CC   protein. This DNA comprises the coding sequence for amino acid residues 1
CC   -881 of human tankyrase. Tankyrase is involved in regulation of telomere
CC   length through TRF1 and may directly modulate the effect of TRF1. This
CC   gene is used to produce recombinant tankyrase, or its fragments, and
CC   these are used to identify specific modulators which are potential drugs
CC   to control telomere length homeostasis for countering telomere shortening
CC   associated with aging (e.g. atrophy of skin, macular degeneration or
CC   atherosclerosis) or diseases such as ataxia telangiectasia. This can also
CC   be used to identify homologous genes and in gene therapy (including
CC   antisense or ribozyme expression). Transgenic animals having at least one
CC   disrupted endogenous allele for tankyrase are used in drug screening
XX
SQ   Sequence 4657 BP; 1268 A; 1162 C; 1160 G; 1067 T; 0 U; 0 Other;

  Query Match             35.9%;  Score 1369.4;  DB 3;  Length 4657;
  Best Local Similarity   63.1%;  Pred. No. 0;

  Matches 2501;  Conservative    0;  Mismatches  956;  Indels  508;  Gaps    5;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||   |  |  | | | || |   || |||      ||||   || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              |||||  || ||||| |||||||||    || || ||||||||||||     | |  || ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||      |||| ||| ||| |||| ||  | || ||||||||||| ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              ||||  ||  ||  ||  ||| |  ||  ||  |||| |||||  |||||||  || |||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              ||  ||  ||  ||  ||  |||||||||||  || |||||||  |||||||  ||  ||  | |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              |  |||  |  ||  ||  ||  |||||  |  |||||  |||||  |||||  |||||  |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||  ||||||||  |||||  |||||  ||||||||||||  |  |||||  |||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              ||||  ||  ||  ||  ||  |||||  |   ||||  ||||  |  ||||||||  ||  ||  |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              ||  ||  |||||  ||||||||  ||||||||  ||||||  |||||   |||  ||||||||  |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              |||||||||  |  ||||||  ||||  ||  ||  |||||||  ||  |||||  ||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              |||||||  |||||  ||||||  |  |||||  ||  ||  ||||||||     ||  ||  |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
              |  ||  ||  |||||  ||||||||  |||||  |||||  |||||||    ||  ||  ||  |  ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              |||||  |||||  |||||  ||||||||||  ||  |||||  |  |  ||||||||  ||  |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

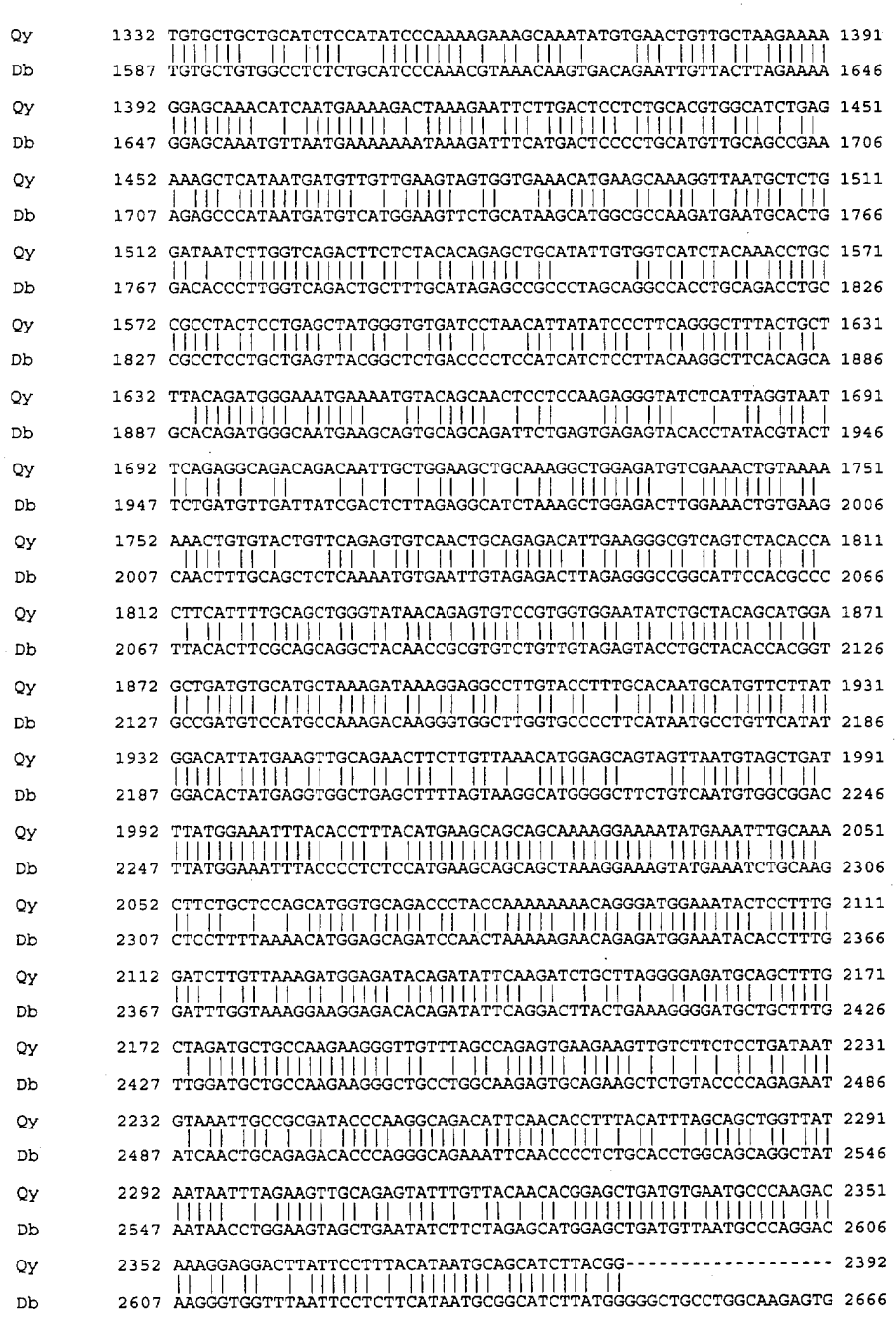
Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              ||  |||||  |||||  |  |||||  ||  |||||  ||  ||  |||||  |||||  ||||  |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

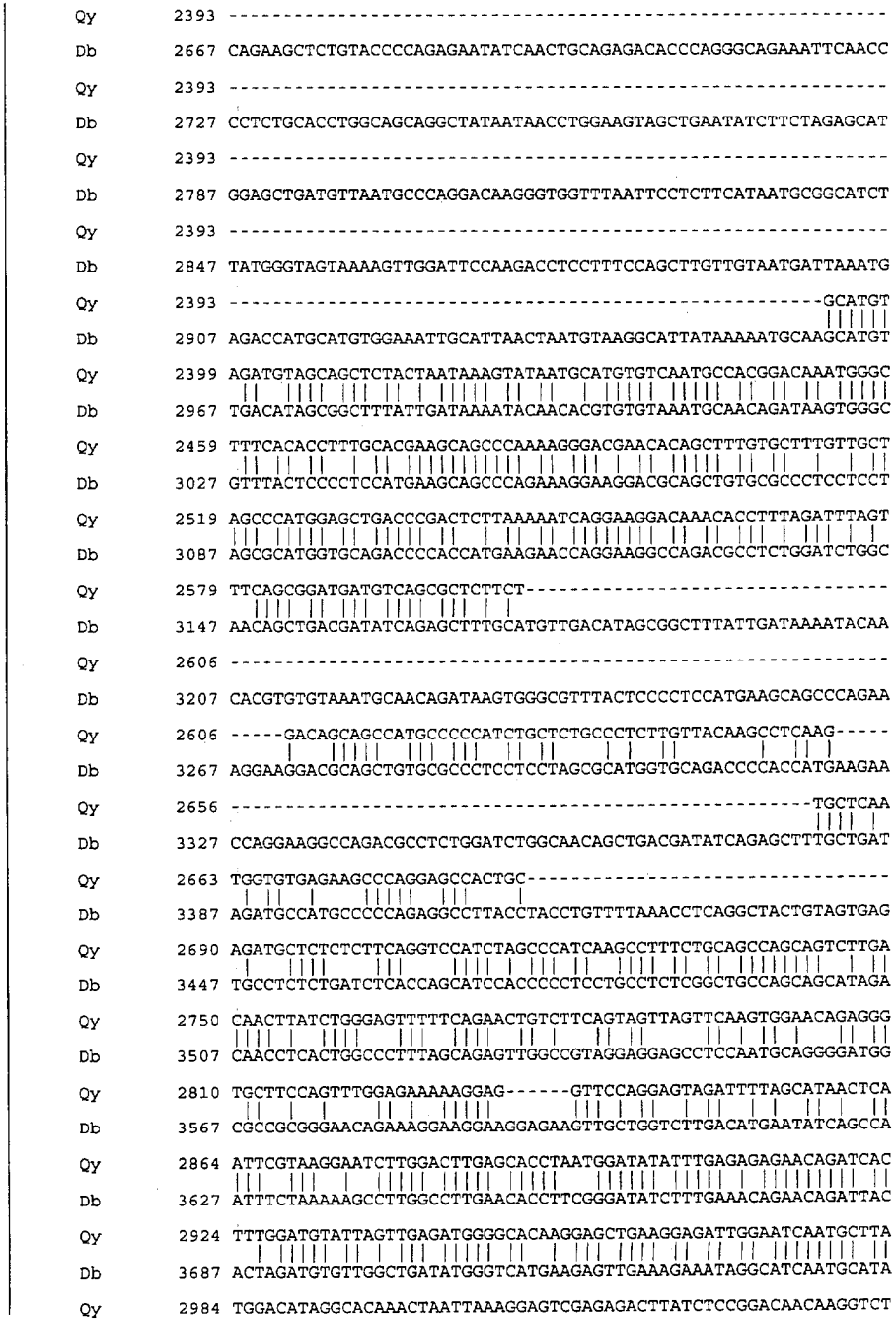
Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              ||  |||||  ||  |||  |  |  ||  ||||  ||  ||  ||  ||  |  ||||  ||  ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
              |||||||||  |  ||  ||||||  ||  ||  ||  |  |  ||  |||||  |  ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              ||||||||  ||  ||  ||  |||||  ||  |||||||  ||  |  ||  |  |  ||||||
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
                 |||  ||||||||||  |  ||||||||  ||  ||  ||  |  ||||||||||||  ||||
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586
```





```
Db       3747 TGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAAGGCAC 3806

Qy       3044 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 3103

Db       3807 CAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTTGCTCC 3866

Qy       3104 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 3163

Db       3867 AGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAG 3926

Qy       3164 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 3223

Db       3927 AGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGT 3986

Qy       3224 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 3283

Db       3987 TGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAA 4046

Qy       3284 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3343

Db       4047 TCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTAT 4106

Qy       3344 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 3403

Db       4107 TCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTA 4166

Qy       3404 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 3463

Db       4167 TTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGG 4226

Qy       3464 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 3523

Db       4227 CTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTCTGTAG 4286

Qy       3524 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 3583

Db       4287 AGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCC 4346

Qy       3584 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 3643

Db       4347 AGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGT 4406

Qy       3644 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 3703

Db       4407 CATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATGAAGCC 4466

Qy       3704 TGAAG 3708

Db       4467 AGAAG 4471
```

RESULT 32
ABK48635
ID ABK48635 standard; cDNA; 4657 BP.
XX
AC ABK48635;
XX
DT 12-AUG-2002 (first entry)
XX
DE Truncated human tankyrase #2 cDNA.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; ss; gene; cytostatic;
KW ADP ribosylation; telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   6. .2855
FT   /*tag= a
FT   /product= "Truncated tankyrase #2"

XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR P-PSDB; AAU79539.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 141-145; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC sequence presented is the truncated human tankyrase #2 cDNA. Note:
CC Nucleotides 1-2648 are the same as the wild-type human tankyrase sequence
CC (ABK48629)
XX
SQ Sequence 4657 BP; 1268 A; 1162 C; 1160 G; 1067 T; 0 U; 0 Other;

Query Match 35.9%; Score 1369.4; DB 6; Length 4657;
Best Local Similarity 63.1%; Pred. No. 0;
Matches 2501; Conservative 0; Mismatches 956; Indels 508; Gaps 5;

```
Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311

Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371

Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431

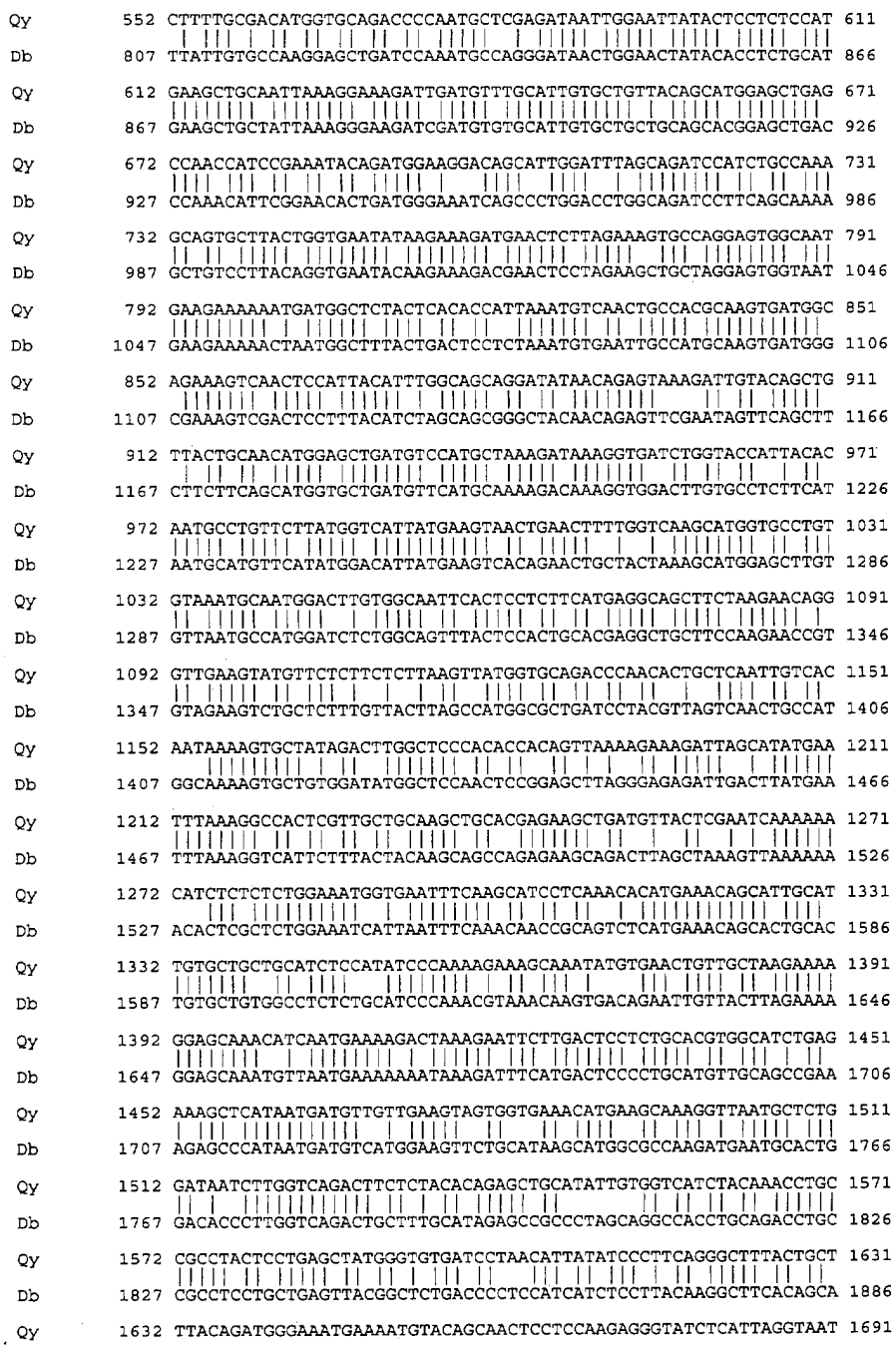
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

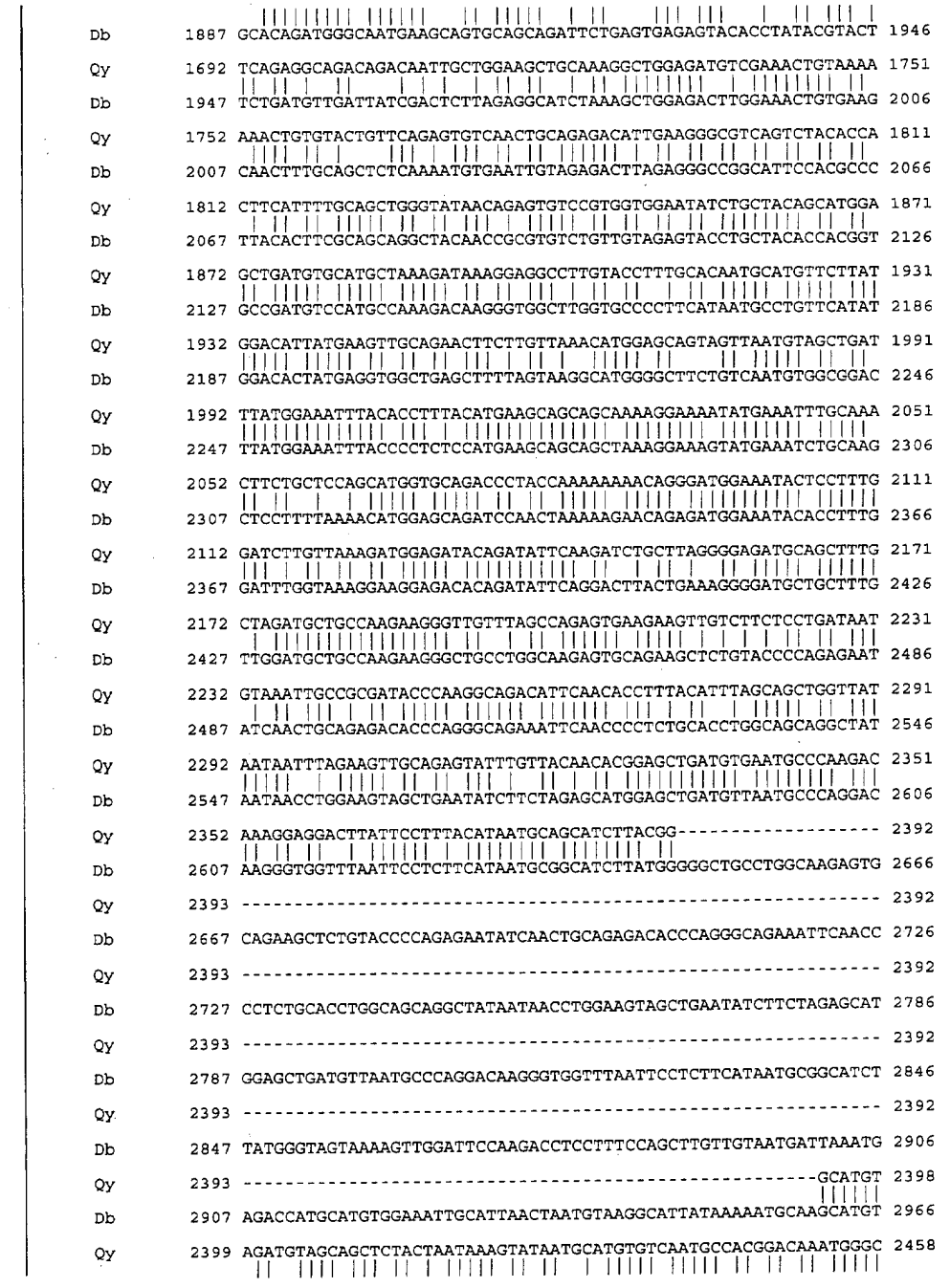
Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491

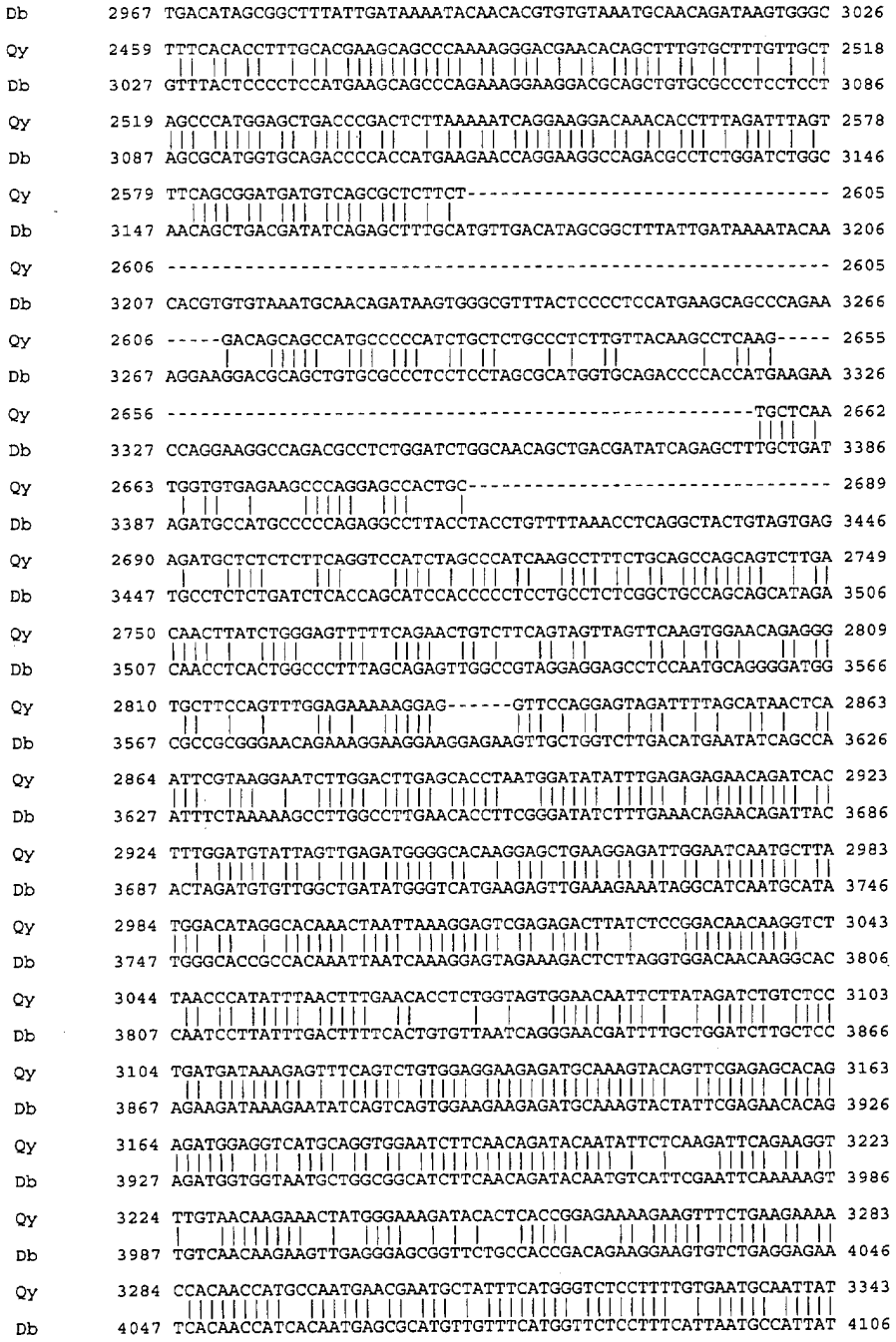
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551

Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806
```









RESULT 33
ABA08564
ID ABA08564 standard; cDNA; 1466 BP.
XX
AC ABA08564;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human Grb7 effector protein homologue-encoding cDNA, SEQ ID NO:340.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11320.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating

PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 1; Page 453; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention
XX
SQ Sequence 1466 BP; 468 A; 287 C; 321 G; 390 T; 0 U; 0 Other;

Query Match 32.5%; Score 1240.4; DB 4; Length 1466;
Best Local Similarity 99.1%; Pred. No. 7.9e-311;
Matches 1247; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy       1042 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 1101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 60

Qy       1102 GTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 1161
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 GTTCTCTTCTCTTAAGTTATGGGGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 120

Qy       1162 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 1221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 180

Qy       1222 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 1281
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 240

Qy       1282 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 1341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 300

Qy       1342 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 1401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 360

Qy       1402 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 1461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 420

Qy       1462 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 1521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 480

Qy       1522 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 1581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 540

Qy       1582 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 1641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 600

Qy       1642 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 1701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 660

Qy       1702 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTA 1761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTA 720

Qy       1762 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 1821
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 780

Qy       1822 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 1881
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 840

Qy       1882 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 1941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 900

Qy       1942 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 2001
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 960

Qy       2002 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 2061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 1020

Qy       2062 AGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 2121
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 1080

Qy       2122 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 2181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 1140

Qy       2182 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 2241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 1200

Qy       2242 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGGTAAGTGAATGAAGTT 1258
```

RESULT 34
AAK99406
ID AAK99406 standard; DNA; 2409 BP.
XX
AC AAK99406;
XX
DT 27-JUN-2002 (first entry)
XX
DE DNA of APP related human homologue hCP33787.
XX
KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW amyloid precursor protein; tissue-specific expression control; human APP;

```
KW   APP pathway modulator; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2409
FT                   /*tag=  a
FT                   /product= "Protein of human homologue hCP33787"
XX
PN   WO200226820-A2.
XX
PD   04-APR-2002.
XX
PF   01-OCT-2001; 2001WO-EP011345.
XX
PR   29-SEP-2000; 2000US-0236893P.
PR   14-JUN-2001; 2001US-0298309P.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   Cohen D,  Dengler UJ,  Finelli AL,  Freuler F,  Konsolaki M;
PI   Reinhardt MWHM,  Zusman S;
XX
DR   WPI; 2002-315796/35.
DR   P-PSDB; AAO20512.
XX
PT   New transgenic fly, containing DNA encoding an Abeta portion of human
PT   APP, useful for identifying agents which modulate the APP pathway and
PT   which can be used to treat Alzheimer's disease.
XX
PS   Example 4; Page 112-113; 129pp; English.
XX
CC   The invention relates to a transgenic fly whose genome comprises DNA
CC   encoding a polypeptide having the Abeta portion of human amyloid
CC   precursor protein (APP), fused to a signal sequence. The DNA sequence
CC   encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC   the specification. The DNA sequence is operably linked to a tissue-
CC   specific expression control sequence. Expression of the sequence gives
CC   the fly an altered phenotype. The purpose of the invention is for
CC   identifying agents that inhibit or promote the expression and/or function
CC   of genes or encoded polypeptides which modify the APP pathway. The agent
CC   is a compound, triple helix DNA, antisense oligonucleotide, double
CC   stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC   to treat conditions such as Alzheimer's disease. The agent can be used as
CC   an APP pathway modulator or in gene therapy. This polynucleotide sequence
CC   represents the DNA of the APP related human homologue hCP33787
XX
SQ   Sequence 2409 BP; 714 A; 480 C; 593 G; 622 T; 0 U; 0 Other;

  Query Match           31.3%;  Score 1192.6;  DB 6;  Length 2409;
  Best Local Similarity  67.4%;  Pred. No. 2.6e-298;
  Matches 2372;  Conservative    0;  Mismatches   19;  Indels 1128;  Gaps    4;

Qy        225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy        285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy        345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy        405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy        465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy        525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy        585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

Qy        645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy        705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATAT------------ 752
              |||||||||||||||||||||||||||||||||||||||||||| | | |
Db        481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTAAGTCTGTATACTCTGGT 540

Qy        753 ------AAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATG 806
                    | |||   || |       |  ||||||||||||||||||||||||||||||
Db        541 TATTCCAGGAAGCCTGTAAAGAACAACCTTGCCAGGAGTGGCAATGAAGAAAAAATGATG 600

Qy        807 GCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCA 866
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGA------------ 648

Qy        867 TTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGA 926
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 ------------------------------------------------------------ 648

Qy        927 GCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTAT 986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 ------------------------------------------------------------ 648

Qy        987 GGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGAC 1046
                                                  ||||||||||||||||||||||||||||
Db        649 ----------------------------AAGCATGGTGCCTGTGTAAATGCAATGGAC 678

Qy       1047 TTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCT 1106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCT 738

Qy       1107 CTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATA 1166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 CTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATA 798

Qy       1167 GACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCG 1226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        799 GACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCG 858

Qy       1227 TTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCTCTGGAA 1286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        859 TTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCTCTGGAA 918

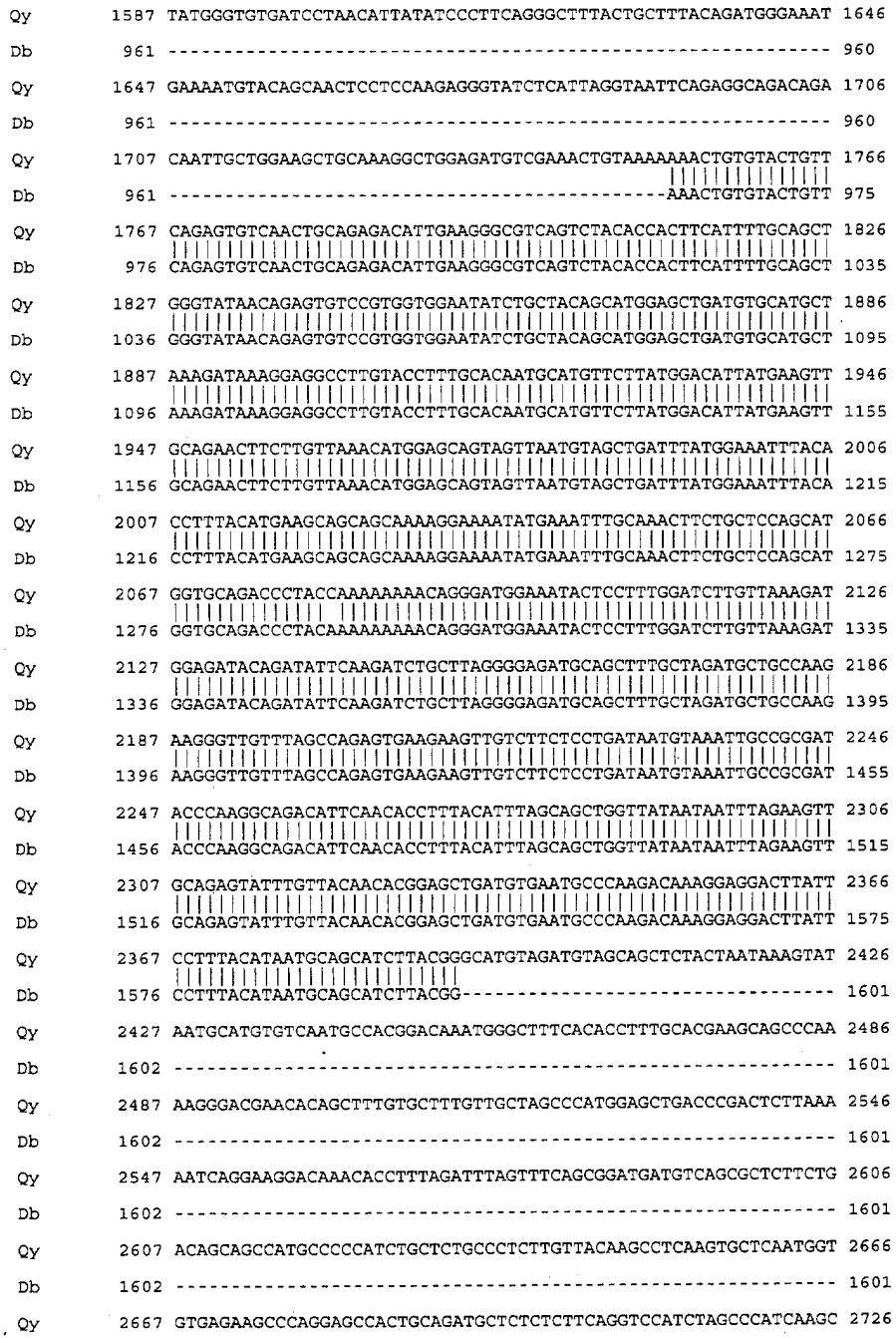
Qy       1287 ATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCT 1346
              ||||||||||||||||||||||||||||||||||||||||
Db        919 ATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTG---------------------- 960

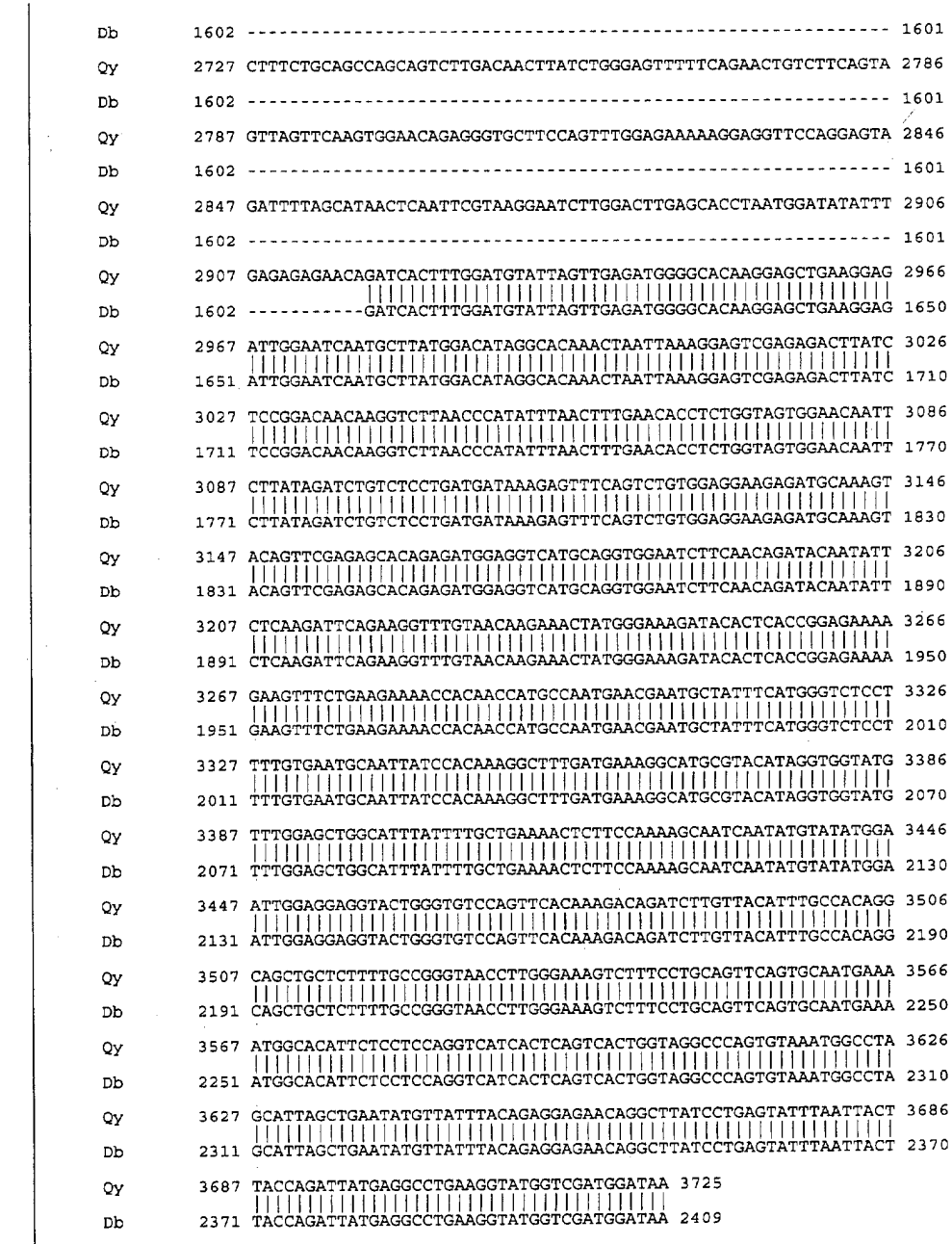
Qy       1347 CCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAAT 1406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ------------------------------------------------------------ 960

Qy       1407 GAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGAT 1466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ------------------------------------------------------------ 960

Qy       1467 GTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAG 1526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ------------------------------------------------------------ 960

Qy       1527 ACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGC 1586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ------------------------------------------------------------ 960
```





RESULT 35

```
AAC89701
ID   AAC89701 standard; cDNA; 1242 BP.
XX
AC   AAC89701;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Human adult T-cell leukaemia cDNA clone HOM-TALL1-5(5').
XX
KW   Human; cytostatic; gene therapy; vaccine; breast cancer;
KW   T cell leukaemia; cancer associated antigen; ss.
XX
OS   Homo sapiens.
XX
PN   WO200072021-A2.
XX
PD   30-NOV-2000.
XX
PF   24-MAY-2000; 2000WO-US014391.
XX
PR   26-MAY-1999;   99US-00320092.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Kuimov A,  Lagarkova M,  Koroleva E,  Turetskaya R,  Vdovichenko K;
PI   Mescheryakov A,  Litchinitser M,  Kuprash D,  Nedospasov S,  Tureci O;
PI   Sahin U,  Pfruendschuh M,  Old LJ,  Knuth A,  Jager E;
XX
DR   WPI; 2001-032067/04.
DR   P-PSDB; AAB48571.
XX
PT   Diagnosing disorder characterized by expression of human cancer
PT   associated antigen precursor, involves detecting interaction of an agent
PT   with novel NA Group 1 nucleic acid molecule encoding the antigen
PT   precursor.
XX
PS   Example 1; Page 74-75; 87pp; English.
XX
CC   The present sequence is given in a specification relating to nucleic
CC   acids and encoded polypeptides which are cancer associated antigens
CC   expressed in patients afflicted with breast cancer and/or T cell
CC   leukaemia. A disorder characterised by expression of human cancer
CC   associated antigen may be diagnosed by contacting a biological sample
CC   with an agent that specifically binds to the nucleic acid that encodes
CC   the antigen, complexed with human leukocyte antigen molecule, and
CC   determining the interaction between the agent and the nucleic acid.
CC   Cancer associated antigens, the nucleotides encoding them, antibodies
CC   against them and the pharmaceutical compositions comprising them are
CC   useful for diagnosing, monitoring and treating the diseases characterised
CC   by the expression of one or more cancer associated antigens
XX
SQ   Sequence 1242 BP; 391 A; 249 C; 269 G; 328 T; 0 U; 5 Other;

  Query Match           30.7%;  Score 1172.2;  DB 4;  Length 1242;
  Best Local Similarity   97.8%;  Pred. No. 3.5e-293;
  Matches 1218;  Conservative    0;  Mismatches   23;  Indels    5;  Gaps    3;

Qy        529 GTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATA 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATA 60

Qy        589 ATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTG 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTG 120

Qy        649 TGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGG 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGG 180

Qy        709 ATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCT 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCT 240

Qy        769 TAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATG 828
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGAAAGTGCCAGGNGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATG 300

Qy        829 TCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATA 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATA 360

Qy        889 ACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATA 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATA 420

Qy        949 AAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAAC 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAAC 480

Qy       1009 TTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTC 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTC 540

Qy       1069 ATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAG 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAG 600

Qy       1129 ACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGT 1188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGT 660

Qy       1189 TAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAG 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAG 720

Qy       1249 CTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTC 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTC 780

Qy       1309 AAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAA 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAA 840

Qy       1369 TATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGA 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGA 900

Qy       1429 CTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAAC 1488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAAC 960

Qy       1489 ATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAT 1548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAT 1020

Qy       1549 ATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTA 1608
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTGTGGTCATCTACACACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTA 1080

Qy       1609 TATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCC 1668
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAA--TGAAATGTACAGCAACTNCTNC 1138

Qy       1669 AAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA-TTGCTGGAAGCTGCAAAG 1727
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 AAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACANTTTNCTGGTAAGCTGCAAA 1198

Qy       1728 GCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTG 1773
              |   |||||||||||      |||||||||||||||||||||||||
Db       1199 GGCTGAGATGTCGAA--CTCTAAAAAACTGTGTACTGTTCAGAGTG 1242


RESULT 36
AAF63990
ID   AAF63990 standard; DNA; 3308 BP.
XX
```

```
AC   AAF63990;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 expression plasmid coding sequence SEQ ID NO: 177.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66301.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 7; Page 232-236; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 3308 BP; 958 A; 730 C; 900 G; 720 T; 0 U; 0 Other;

  Query Match           30.6%;  Score 1168.4;  DB 5;  Length 3308;
  Best Local Similarity   99.9%;  Pred. No. 6e-292;
  Matches 1169;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2553 GAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCA 2612
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       2101 GAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCA 2160

Qy       2613 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2672
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2220

Qy       2673 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2732
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2280

Qy       2733 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2792
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2340

Qy       2793 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2852
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2400

Qy       2853 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2912
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2460

Qy       2913 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2972
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       2461 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2520

Qy       2973 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 3032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 2580

Qy       3033 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 3092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 2640

Qy       3093 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 3152
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 2700

Qy       3153 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 3212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 2760

Qy       3213 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 3272
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 2820

Qy       3273 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 3332
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 2880

Qy       3333 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 3392
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 2940

Qy       3393 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3000

Qy       3453 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3060

Qy       3513 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3120

Qy       3573 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3632
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3180

Qy       3633 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3240

Qy       3693 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
             |||||||||||||||||||||||||||||
Db       3241 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3270

RESULT 37
ACA57422
ID   ACA57422 standard; cDNA; 1121 BP.
XX
AC   ACA57422;
XX
DT   10-JUN-2003  (first entry)
XX
DE   Human adipocyte Selected Interacting domain, SID, cDNA #509.
XX
KW   Human; ss; gene; prey; adipocyte; SID; selected interacting domain;
KW   anorectic; antidiabetic; protein-protein interaction; diabetes;
KW   yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS   Homo sapiens.
XX
PN   WO200286122-A2.
XX
```

PD 31-OCT-2002.
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Daviet L;
XX
DR WPI; 2003-103412/09.
DR P-PSDB; ABU70878.
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
PS Claim 7; Page 276; 382pp; English.
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM)) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence encodes a SID
CC (prey) protein of the invention
XX
SQ Sequence 1121 BP; 354 A; 228 C; 241 G; 298 T; 0 U; 0 Other;

Query Match 29.4%; Score 1121; DB 7; Length 1121;
Best Local Similarity 100.0%; Pred. No. 6.4e-280;
Matches 1121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       458 TCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGC 517
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 TCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGC 60

Qy       518 ATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAA 577
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAA 120

Qy       578 TGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGA 637
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGA 180

Qy       638 TGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAG 697
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAG 240

Qy       698 GACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAA 757
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAA 300

Qy       758 AGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCAC 817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCAC 360

Qy       818 ACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGC 877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGC 420

Qy       878 AGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCA 937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCA 480

Qy       938 TGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGA 997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGA 540

Qy       998 AGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATT 1057
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 AGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATT 600

Qy      1058 CACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAG 1117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAG 660

Qy      1118 TTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCC 1177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 TTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCC 720

Qy      1178 CACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGC 1237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGC 780

Qy      1238 TGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTT 1297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTT 840

Qy      1298 CAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAA 1357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 CAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAA 900

Qy      1358 AAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAA 1417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 AAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAA 960

Qy      1418 AGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGT 1477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 AGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGT 1020

Qy      1478 AGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACA 1537
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACA 1080

Qy      1538 CAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTAC 1578
            |||||||||||||||||||||||||||||||||||||||||
Db      1081 CAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTAC 1121
```

RESULT 38
AAF63973
ID AAF63973 standard; DNA; 1103 BP.
XX
AC AAF63973;
XX
DT 05-APR-2001 (first entry)
XX
DE Human TRF1 TANK2 binding domain coding sequence SEQ ID NO: 159.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX

```
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66300.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 6; Page 226-228; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1103 BP; 352 A; 219 C; 239 G; 293 T; 0 U; 0 Other;

  Query Match           28.4%;  Score 1084.8;  DB 5;  Length 1103;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-270;
  Matches 1086;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1056 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 1115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 68

Qy       1116 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 1175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 128

Qy       1176 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 1235
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 188

Qy       1236 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 1295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 248

Qy       1296 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 1355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 308

Qy       1356 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 368

Qy       1416 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 428

Qy       1476 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 488

Qy       1536 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 1595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 548

Qy       1596 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 1655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 608

Qy       1656 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 1715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 668

Qy       1716 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTC 1775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTC 728

Qy       1776 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 1835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 788

Qy       1836 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 1895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 848

Qy       1896 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 1955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 908

Qy       1956 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 2015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 968

Qy       2016 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 2075
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 1028

Qy       2076 CCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 2135
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 CCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 1088

Qy       2136 GATATTCA 2143
             |||||| |
Db       1089 GATATTTA 1096

RESULT 39
AAH16624
ID   AAH16624 standard; cDNA; 3006 BP.
XX
AC   AAH16624;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:15732.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
```

```
XX
PS   Claim 8; SEQ ID NO 15732; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 3006 BP; 892 A; 505 C; 584 G; 1025 T; 0 U; 0 Other;

  Query Match            28.2%;  Score 1076.8;  DB 4;  Length 3006;
  Best Local Similarity  99.8%;  Pred. No. 3.3e-268;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2734 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 2793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 60

Qy       2794 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 2853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 120

Qy       2854 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 2913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 180

Qy       2914 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 2973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 240

Qy       2974 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 3033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 300

Qy       3034 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 3093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 360

Qy       3094 ATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 3153
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        361 ATCTGTCTCCTGATGATAAAGAGTTTCCGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 420

Qy       3154 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 3213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 480

Qy       3214 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 540

Qy       3274 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 600

Qy       3334 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 660

Qy       3394 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3453
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 720

Qy       3454 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3513
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 780

Qy       3514 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 3573
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 840

Qy       3574 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3633
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 900

Qy       3634 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3693
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        901 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTAAATTACTTACCAGA 960

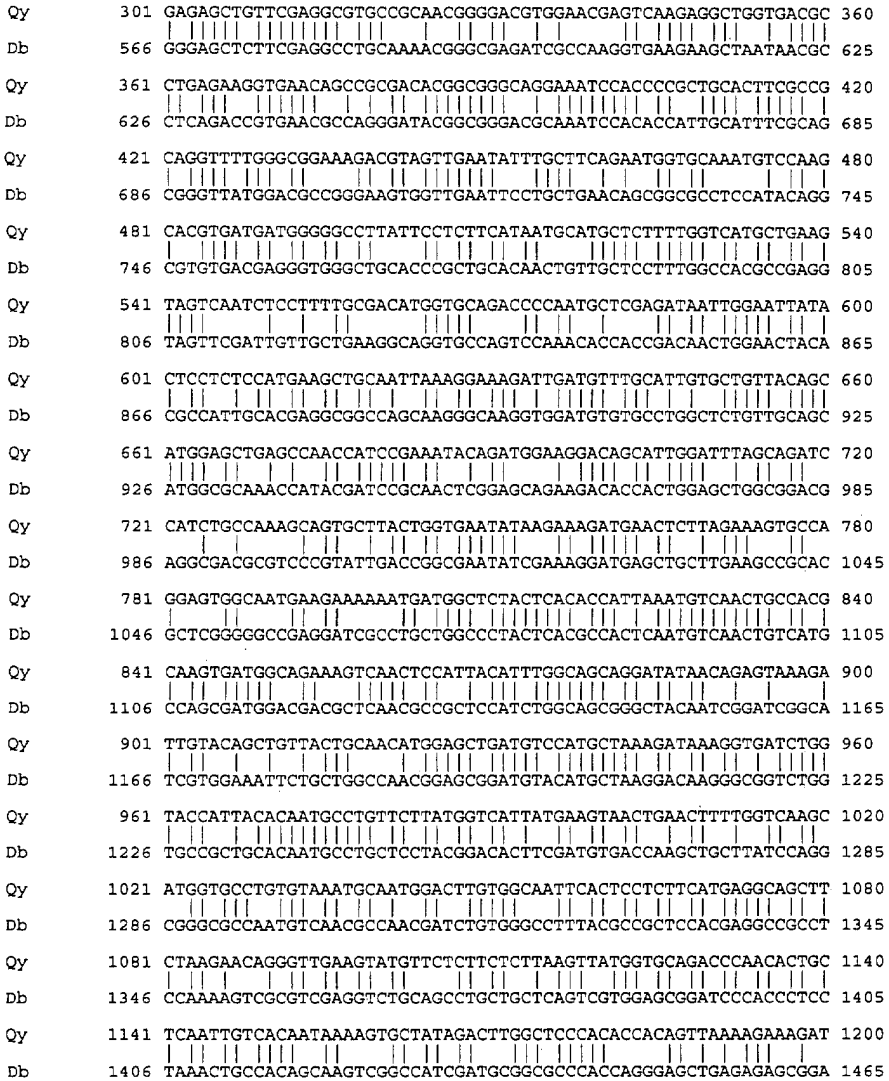
Qy       3694 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 1020

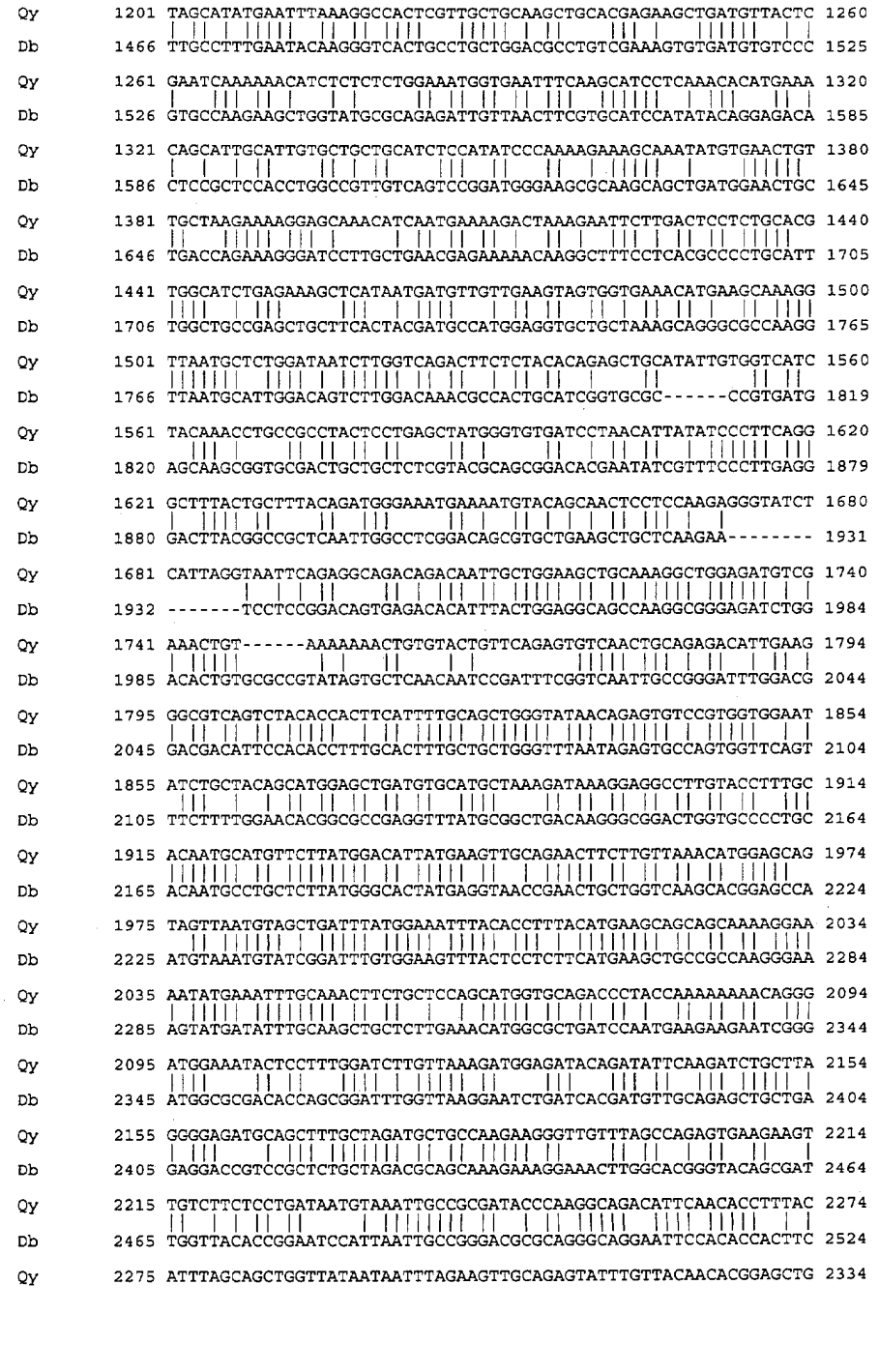
Qy       3754 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 1080

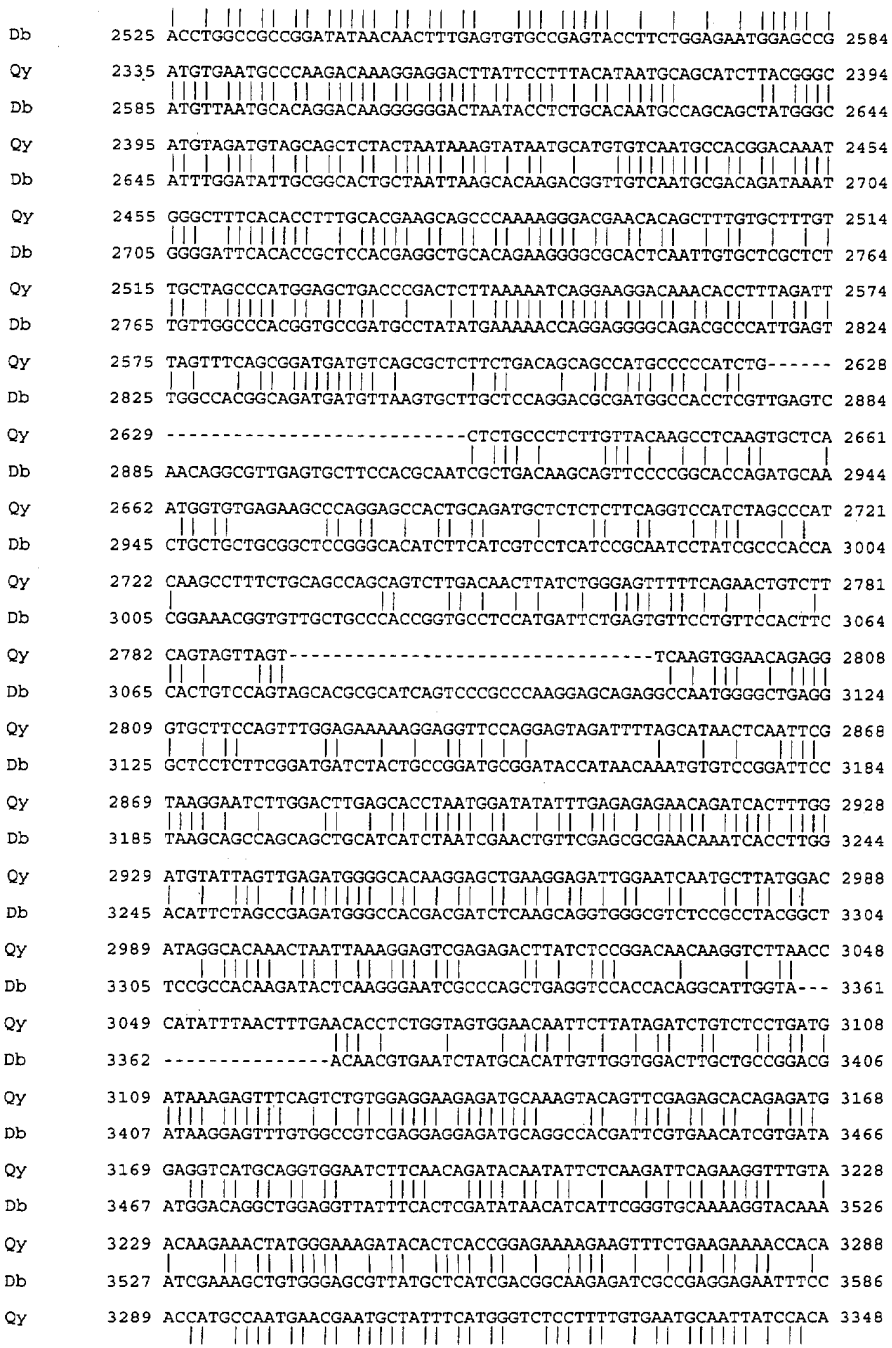
RESULT 40
ABL08155
ID   ABL08155 standard; cDNA; 5484 BP.
XX
AC   ABL08155;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 18947.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB64052.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 18947; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
```

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5484 BP; 1346 A; 1467 C; 1495 G; 1176 T; 0 U; 0 Other;

Query Match 28.0%; Score 1068.4; DB 4; Length 5484;
Best Local Similarity 58.6%; Pred. No. 7e-266;
Matches 2038; Conservative 0; Mismatches 1331; Indels 111; Gaps 6;

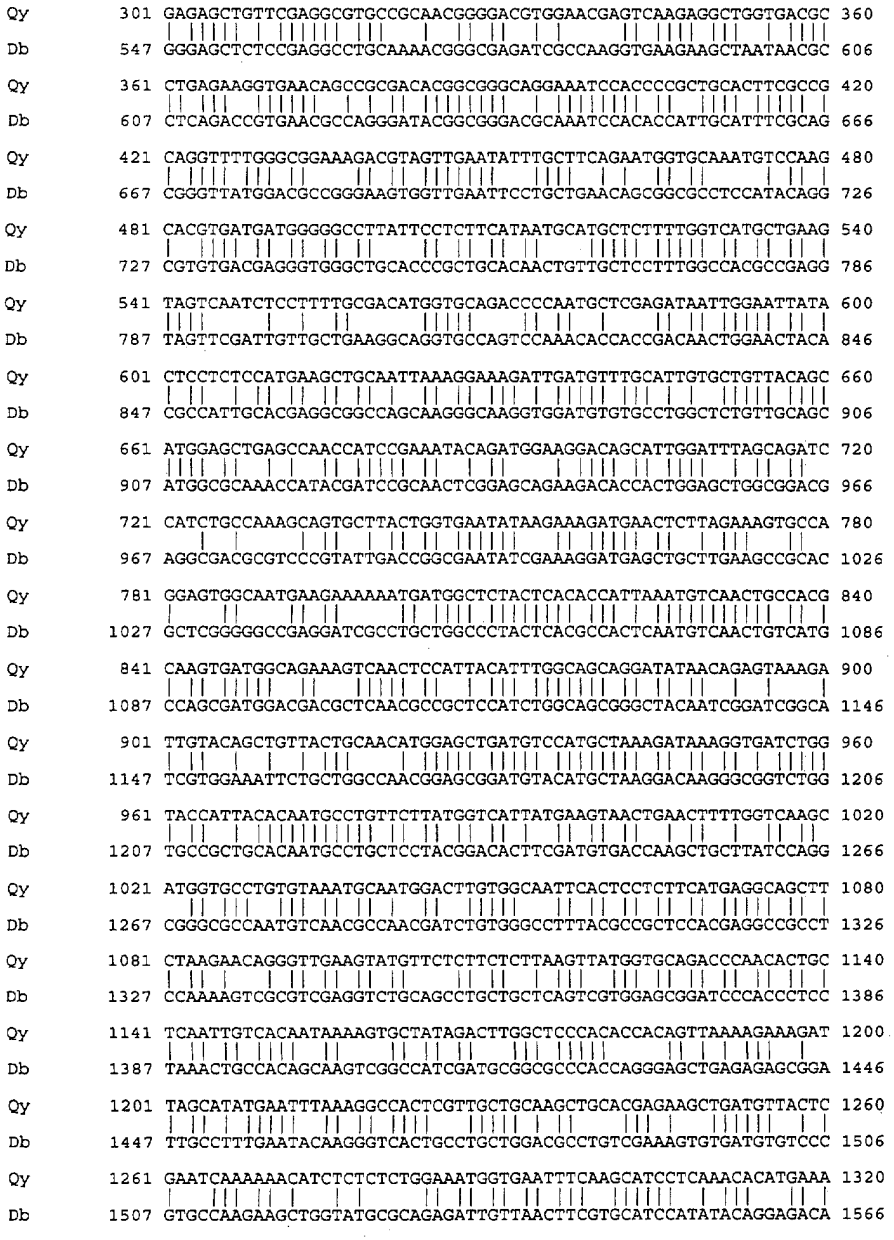


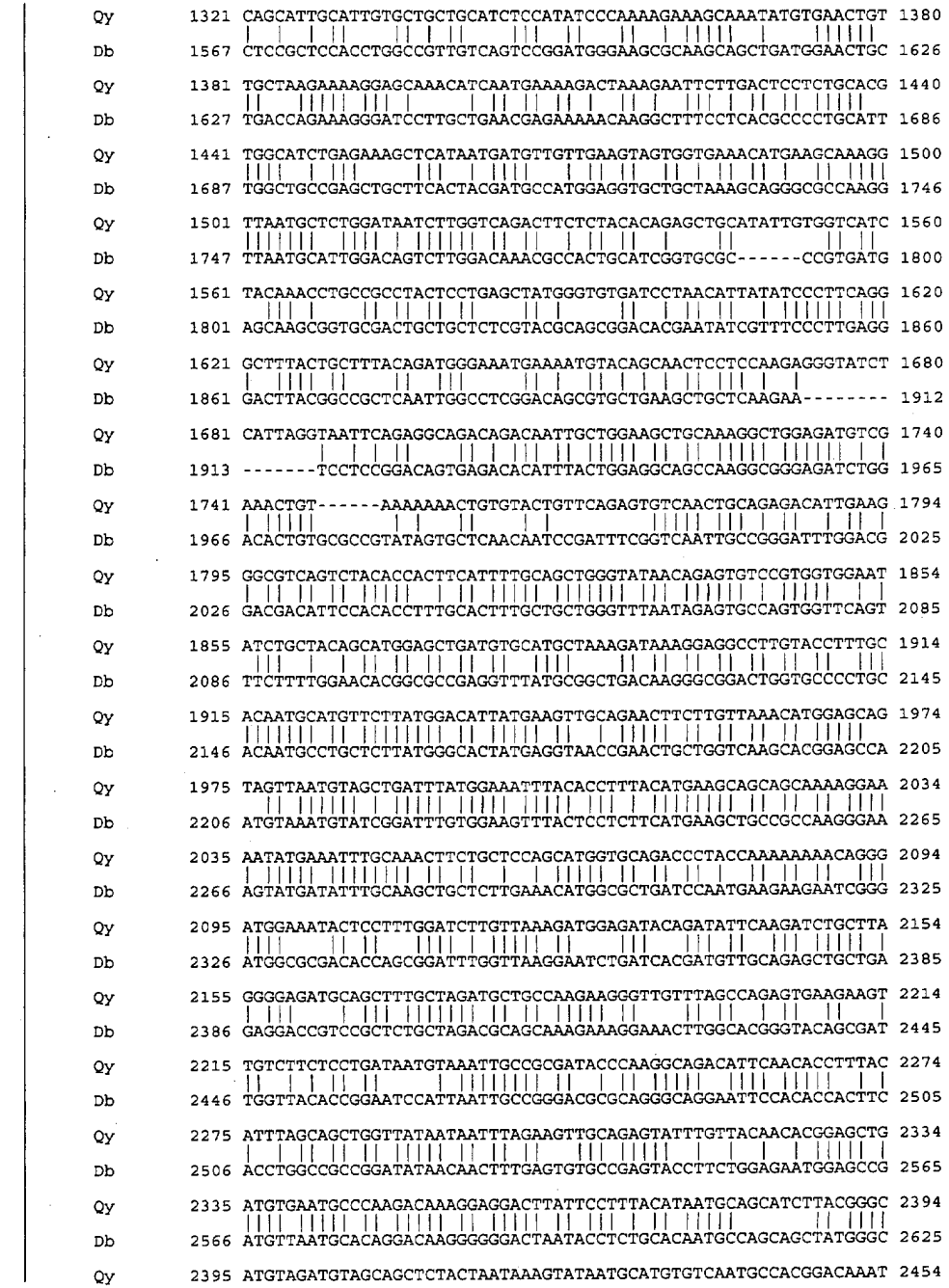


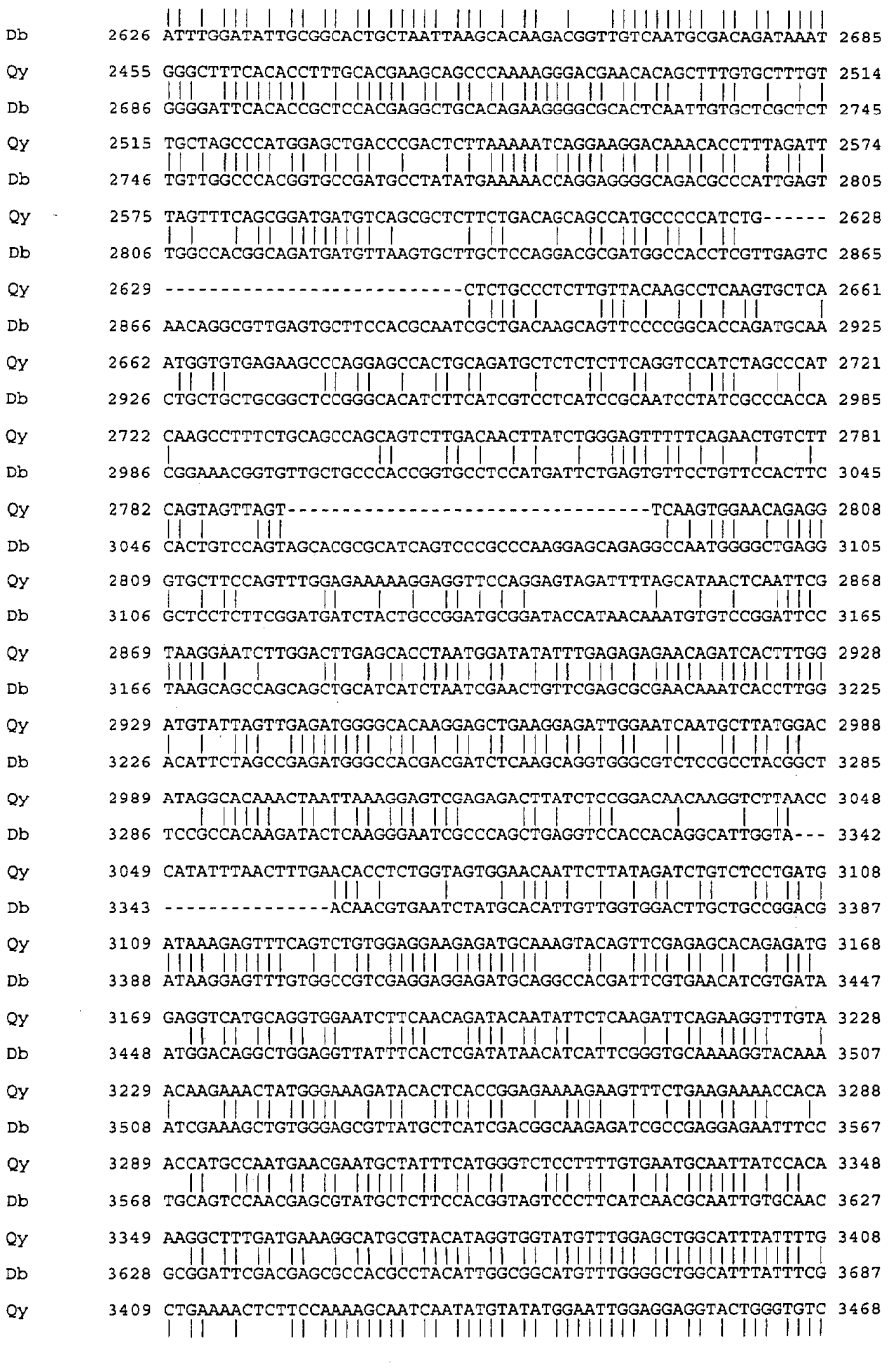


RESULT 41
AAF63955
ID AAF63955 standard; DNA; 5482 BP.
XX
AC AAF63955;
XX
DT 05-APR-2001 (first entry)
XX
DE Drosophila tankyrase gene SEQ ID NO: 138.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66297.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 2; Page 210-215; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 5482 BP; 1359 A; 1461 C; 1493 G; 1169 T; 0 U; 0 Other;

Query Match 28.0%; Score 1066.8; DB 5; Length 5482;
Best Local Similarity 58.5%; Pred. No. 1.8e-265;
Matches 2037; Conservative 0; Mismatches 1332; Indels 111; Gaps 6;







RESULT 42
AAC89703
ID AAC89703 standard; cDNA; 1007 BP.
XX
AC AAC89703;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human adult T-cell leukaemia cDNA clone MO-BC-203(5').
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR P-PSDB; AAB48572.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 1; Page 78; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 1007 BP; 315 A; 200 C; 219 G; 273 T; 0 U; 0 Other;

```
  Query Match           25.9%;  Score 990;  DB 4;  Length 1007;
  Best Local Similarity   99.0%;  Pred. No. 5.9e-246;
  Matches  996;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 60

Qy       1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 120

Qy       1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 180

Qy       1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240

Qy       1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 300

Qy       1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 360

Qy       1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
              |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db        361 AAAGAAAGCAAATATGTGAACTGTTGCTAAAAAAAGGATCAAACATCAATGAAAAGACTA 420

Qy       1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480

Qy       1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540

Qy       1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 600

Qy       1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 660

Qy       1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 720

Qy       1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        721 AAGCTGCAAAGGCTGGAGATGTCAAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 780

Qy       1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 840

Qy       1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        841 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGAGAAAG 900

Qy       1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 960

Qy       1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATT 2002
              |||||||||||||||||| | |||||  | || ||||||||||
Db        961 TTGTTAAACATGGAGCAGTATTCAATGTTACCGAATTATGGAAATT 1006
```

```
RESULT 43
ADE53994
ID   ADE53994 standard; cDNA; 797 BP.
XX
AC   ADE53994;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human prostate cancer cDNA #341.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.
XX
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.
PA   (PEAR/) PEARSON C I.
XX
PI   Faris M,  Pearson CI;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 1; SEQ ID NO 341; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
CC   cancer cDNA of the invention.
XX
SQ   Sequence 797 BP; 256 A; 144 C; 184 G; 212 T; 0 U; 1 Other;
```

```
  Query Match           20.8%;  Score 794.4;  DB 9;  Length 797;
  Best Local Similarity   99.7%;  Pred. No. 2.9e-195;
  Matches  795;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1363 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 60

Qy       1423 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 120

Qy       1483 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 180

Qy       1543 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 240
```

```
Qy       1603 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 1662
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACATTATATCCCTTCANGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 300

Qy       1663 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 360

Qy       1723 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 420

Qy       1783 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 480

Qy       1843 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 540

Qy       1903 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 600

Qy       1963 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 660

Qy       2023 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2082
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        661 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 720

Qy       2083 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 780

Qy       2143 AAGATCTGCTTAGGGGA 2159
             ||||||||||||||||
Db        781 AAGATCTGCTTAGGGGA 797
```

RESULT 44
AAC77578
ID AAC77578 standard; cDNA; 1344 BP.
XX
AC AAC77578;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF3133 polynucleotide sequence SEQ ID NO:6265.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB43369.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.
XX
PS Claim 5; Page 5454; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 1344 BP; 418 A; 243 C; 259 G; 422 T; 0 U; 2 Other;

Query Match 20.6%; Score 786; DB 3; Length 1344;
Best Local Similarity 100.0%; Pred. No. 5.8e-193;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3028 CCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTC 3087
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 CCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTC 85

Qy       3088 TTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTA 3147
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 TTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTA 145

Qy       3148 CAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTC 3207
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 CAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTC 205

Qy       3208 TCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAG 3267
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 TCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAG 265

Qy       3268 AAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTT 3327
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 AAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTT 325

Qy       3328 TTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGT 3387
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 TTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGT 385

Qy       3388 TTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAA 3447
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 TTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAA 445
```

```
Qy       3448 TTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGC 3507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 TTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGC 505

Qy       3508 AGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAA 3567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 AGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAA 565

Qy       3568 TGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAG 3627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAG 625

Qy       3628 CATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTT 3687
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTT 685

Qy       3688 ACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAAT 3747
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 ACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAAT 745

Qy       3748 TCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAA 3807
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 TCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAA 805

Qy       3808 AAAAAA 3813
              ||||||
Db        806 AAAAAA 811
```

RESULT 45
AAH08231
ID AAH08231 standard; cDNA; 838 BP.
XX
AC AAH08231;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:5066.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 1; SEQ ID NO 5066; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 838 BP; 260 A; 143 C; 203 G; 227 T; 0 U; 5 Other;

Query Match 20.6%; Score 784.2; DB 4; Length 838;
Best Local Similarity 98.0%; Pred. No. 1.3e-192;
Matches 823; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

```
Qy       2734 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 2793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 60

Qy       2794 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 2853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 120

Qy       2854 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 2913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 180

Qy       2914 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 2973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 240

Qy       2974 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 3033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 300

Qy       3034 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 3093
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 360

Qy       3094 ATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 3153
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        361 ATCTGTCTCCTGATGATAAAGAGTTTCCGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 420

Qy       3154 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 3213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 480

Qy       3214 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 540

Qy       3274 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 600

Qy       3334 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 660

Qy       3394 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3453
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        661 CTGGCATTTATTTTGCTGAAAACTCTT-CAAAAGCAATCAATATGTATATGGAATTGGAN 719
```

```
Qy       3454 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3513
             || || |||||||||||||||||||||||||| ||||||| |||||||||||| ||||||||
Db        720 GANGT-CTGGGTGTCCAGTTCACAAAGACCGATCTTGGTACATTTGCCACANGCAGCTGC 778

Qy       3514 TCTTTTG--CCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 3571
             ||||||| | ||  | ||||||||||||||||| |||||||||| ||||||| |||||||
Db        779 TCTTTTGGCCGGGTAACCCTTGGGAAAGTCTTTTCTGCAGTTCANTGCAATGNAAATGGC 838
```

Search completed: March 29, 2004, 20:39:00
Job time : 1441.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:44:37 ; Search time 49.812 Seconds
(without alignments)
6513.709 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6375.5	98.6	1262	14	US-10-199-937-107	Sequence 107, App
2	6375.5	98.6	1385	14	US-10-199-937-133	Sequence 133, App
3	6121	94.7	1169	14	US-10-199-937-2	Sequence 2, Appli
4	6121	94.7	1169	14	US-10-199-937-101	Sequence 101, App
5	6108	94.5	1166	10	US-09-972-115A-6	Sequence 6, Appli
6	6108	94.5	1166	14	US-10-163-587A-15	Sequence 15, Appl
7	6108	94.5	1166	14	US-10-199-937-135	Sequence 135, App
8	5877	90.9	1227	10	US-09-849-602-26	Sequence 26, Appl
9	5602	86.7	1074	9	US-09-509-196A-2	Sequence 2, Appli
10	5452	84.3	1267	10	US-09-972-115A-4	Sequence 4, Appli
11	5103.5	79.0	1327	9	US-09-841-835-2	Sequence 2, Appli
12	5103.5	79.0	1327	10	US-09-972-115A-8	Sequence 8, Appli
13	5103.5	79.0	1327	14	US-10-199-937-4	Sequence 4, Appli
14	5031.5	77.8	1333	10	US-09-972-115A-2	Sequence 2, Appli
15	4102	63.5	784	14	US-10-199-937-89	Sequence 89, Appl
16	3997	61.8	1181	14	US-10-199-937-139	Sequence 139, App
17	3959	61.2	756	14	US-10-199-937-91	Sequence 91, Appl
18	3938	60.9	802	9	US-09-964-899-41	Sequence 41, Appl
19	3307	51.2	949	9	US-09-841-835-10	Sequence 10, Appl
20	2722	42.1	522	14	US-10-199-937-99	Sequence 99, Appl
21	2165	33.5	415	12	US-10-276-774-1690	Sequence 1690, Ap
22	2132.5	33.0	673	9	US-09-841-835-8	Sequence 8, Appli
23	2056.5	31.8	1099	14	US-10-199-937-178	Sequence 178, App
24	1895	29.3	362	14	US-10-199-937-160	Sequence 160, App
25	888.5	13.7	1762	14	US-10-205-194-117	Sequence 117, App
26	886.5	13.7	215	14	US-10-199-937-105	Sequence 105, App
27	877	13.6	1330	15	US-10-108-260A-3237	Sequence 3237, Ap
28	874	13.5	160	10	US-09-972-115A-20	Sequence 20, Appl
29	874	13.5	160	14	US-10-199-937-46	Sequence 46, Appl
30	815.5	12.6	1724	9	US-09-964-899-43	Sequence 43, Appl
31	772	11.9	153	14	US-10-106-698-6500	Sequence 6500, Ap
32	760.5	11.8	1053	12	US-10-221-278-343	Sequence 343, App
33	760.5	11.8	1053	15	US-10-291-172-343	Sequence 343, App
34	714.5	11.1	919	15	US-10-108-260A-4122	Sequence 4122, Ap
35	650	10.1	119	14	US-10-199-937-42	Sequence 42, Appl
36	650	10.1	119	14	US-10-199-937-44	Sequence 44, Appl
37	639	9.9	118	14	US-10-199-937-8	Sequence 8, Appli
38	633.5	9.8	1619	14	US-10-369-978-4	Sequence 4, Appli
39	563	8.7	1704	14	US-10-369-978-2	Sequence 2, Appli
40	545.5	8.4	1709	14	US-10-369-978-6	Sequence 6, Appli
41	534	8.3	747	15	US-10-094-749-1924	Sequence 1924, Ap
42	528.5	8.2	3913	15	US-10-334-143-45	Sequence 45, Appl
43	478	7.4	765	12	US-10-182-243-56	Sequence 56, Appl
44	472.5	7.3	657	15	US-10-104-047-2529	Sequence 2529, Ap
45	465	7.2	765	14	US-10-128-174-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-199-937-107
; Sequence 107, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 1262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-107

Query Match 98.6%; Score 6375.5; DB 14; Length 1262;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1225; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```
Qy          6 RGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 65
              ||: |    |||| |  ||||||||||||||||||||||||||||||||||||||||||
Db         29 RGSRGAGSPARGAR-GRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 87

Qy         66 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 147

Qy        126 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 185
```

```
Db       148 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 207

Qy       186 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 245

Db       208 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 267

Qy       246 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 305

Db       268 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 327

Qy       306 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 365

Db       328 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 387

Qy       366 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 425

Db       388 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 447

Qy       426 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 485

Db       448 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 507

Qy       486 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 545

Db       508 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 567

Qy       546 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 605

Db       568 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 627

Qy       606 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 665

Db       628 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 687

Qy       666 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 725

Db       688 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 747

Qy       726 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 785

Db       748 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 807

Qy       786 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 845

Db       808 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 867

Qy       846 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 905

Db       868 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 927

Qy       906 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 965

Db       928 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 987

Qy       966 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1025

Db       988 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1047

Qy      1026 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1085

Db      1048 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1107

Qy      1086 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1145

Db      1108 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1167

Qy      1146 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1205

Db      1168 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1227

Qy      1206 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240

Db      1228 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1262
```

RESULT 2
US-10-199-937-133
; Sequence 133, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 133
; LENGTH: 1385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-133

Query Match 98.6%; Score 6375.5; DB 14; Length 1385;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1225; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```
Qy         6 RGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 65

Db       152 RGSRGAGSPARGAR-GRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLAL 210

Qy        66 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 125

Db       211 LLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNS 270

Qy       126 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 185

Db       271 RDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLL 330

Qy       186 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 245

Db       331 RHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAV 390

Qy       246 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 305

Db       391 LTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL 450

Qy       306 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 365

Db       451 QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVE 510

Qy       366 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 425

Db       511 VCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHL 570

Qy       426 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 485

Db       571 SLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKA 630

Qy       486 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 545

Db       631 HNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ 690

Qy       546 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 605

Db       691 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 750

Qy       606 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 665
```

```
Db        751 FAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLW 810

Qy        666 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 KFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLD 870

Qy        726 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKG 930

Qy        786 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 GLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADP 990

Qy        846 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        991 TLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSS 1050

Qy        906 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1051 PSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLM 1110

Qy        966 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1111 DIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGS 1170

Qy       1026 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1171 GTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTH 1230

Qy       1086 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1231 RRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQY 1290

Qy       1146 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1291 VYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSV 1350

Qy       1206 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||
Db       1351 NGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1385
```

RESULT 3
US-10-199-937-2
; Sequence 2, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-2

Query Match 94.7%; Score 6121; DB 14; Length 1169;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         72 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 60

Qy        132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 120

Qy        192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 180

Qy        252 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 240

Qy        312 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 300

Qy        372 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 360

Qy        432 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 420

Qy        492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 480

Qy        552 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 540

Qy        612 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 600

Qy        672 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 660

Qy        732 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 720

Qy        792 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 780

Qy        852 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 840

Qy        912 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 900

Qy        972 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 960

Qy       1032 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1020

Qy       1092 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1080

Qy       1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1211
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1140

Qy       1212 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
                 |||||||||||||||||||||||||||||
Db       1141 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1169


RESULT 4
US-10-199-937-101
; Sequence 101, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
;   LENGTH: 1169
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-101

  Query Match             94.7%;  Score 6121;  DB 14;  Length 1169;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1169;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          72 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 131
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 60

Qy         132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 120

Qy         192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 251
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 180

Qy         252 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 311
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 240

Qy         312 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 371
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 300

Qy         372 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 360

Qy         432 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 491
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 420

Qy         492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 551
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 480

Qy         552 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 611
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 540

Qy         612 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 671
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 600

Qy         672 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 731
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGC 660

Qy         732 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 791
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 720

Qy         792 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 851
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQE 780

Qy         852 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 911
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 840

Qy         912 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 971
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERE 900

Qy         972 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 1031
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 QITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 960

Qy        1032 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1091
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 LSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVS 1020

Qy        1092 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1080

Qy        1152 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1140

Qy        1212 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
               ||||||||||||||||||||||||||||||
Db        1141 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1169


RESULT 5
US-09-972-115A-6
; Sequence 6, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 1166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-6

  Query Match             94.5%;  Score 6108;  DB 10;  Length 1166;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1166;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
Db        301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
Db        421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
Db        481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
Db        541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
Db        601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy        735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
Db        661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
Db        721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

RESULT 6
US-10-163-587A-15
; Sequence 15, Application US/10163587A
; Publication No. US20030096263A1
; GENERAL INFORMATION:
; APPLICANT: Oliveira, Marcos
; TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZING I
; FILE REFERENCE: 50229-306
; CURRENT APPLICATION NUMBER: US/10/163,587A
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/296,110
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-163-587A-15

Query Match 94.5%; Score 6108; DB 14; Length 1166;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy        255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
Db        181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
Db        241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
Db        301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
Db        361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
Db        421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
Db        481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
Db        541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600
```

```
Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy        735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy        795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy        855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy        915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy        975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy       1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy       1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy       1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140

Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166


RESULT 7
US-10-199-937-135
; Sequence 135, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
;   LENGTH: 1166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-135

  Query Match            94.5%;  Score 6108;  DB 14;  Length 1166;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1166;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy         135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy         195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 180

Qy         255 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 314
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 240

Qy         315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 300

Qy         375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 360

Qy         435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 420

Qy         495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 480

Qy         555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 540

Qy         615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600

Qy         675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 660

Qy         735 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 794
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 720

Qy         795 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 854
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 780

Qy         855 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 914
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 840

Qy         915 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 974
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 900

Qy         975 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 1034
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 960

Qy        1035 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1094
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 1020

Qy        1095 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1080

Qy        1155 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1140
```

```
Qy       1215 IYRGEQAYPEYLITYQIMRPEGMVDG 1240
             |||||||||||||||||||||||||||
Db       1141 IYRGEQAYPEYLITYQIMRPEGMVDG 1166
```

RESULT 8
US-09-849-602-26
; Sequence 26, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-26

Query Match 90.9%; Score 5877; DB 10; Length 1227;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1148; Conservative 8; Mismatches 47; Indels 50; Gaps 5;

```
Qy         12 QGAQRGARVGAA-----------HGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLL 59
              |||||||||||||           |  |: | :|     |  | :|  :|  :||
Db          1 QGAQRGARVGAAMGLRRSGDSREPSGPGPERVFSGGPRPPARGAGAPAPVAGAVAG---- 56

Qy         60 LRLLALLLAVAAARIMSGRRCAGG-----------GAACASAAAEAVEPAARELFEACRN 108
                                   | ||           |:    ||||||||||||||||||
Db         57 --------------------CGGGQDHVGSPLRRRGSGLRDAAAEAVEPAARELFEACRN 96

Qy        109 GDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIP 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 GDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIP 156

Qy        169 LHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 LHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 216

Qy        229 TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP 276

Qy        289 LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD 336

Qy        349 LWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 LWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS 396

Qy        409 LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN 456

Qy        469 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 516

Qy        529 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 576

Qy        589 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 636

Qy        649 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD 696

Qy        709 GDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 GDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 756

Qy        769 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQ 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQ 816

Qy        829 KGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 KGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG 876

Qy        889 VRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGV 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 VRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGV 936

Qy        949 DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLI 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLI 996

Qy       1009 SGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNI 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 SGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNI 1056

Qy       1069 LKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGM 1128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 LKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGM 1116

Qy       1129 FGAGIYFAENSSKSNQYVYGIGGGTGCPVHK-DRSCYICHRQLLFCRVTLGKSFLQFSAM 1187
              |||||||||||||||||||||||||||||     | |   :     ||     ||| |
Db       1117 FGAGIYFAENSSKSNQYVYGIGGGTGVQFTKTDLVTFATAAALL--PGNLGKVFPAVQCN 1174

Qy       1188 KMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              :  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1175 ENGTSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1227
```

RESULT 9
US-09-509-196A-2
; Sequence 2, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
; APPLICANT: DALY, Roger J.
; APPLICANT: SUTHERLAND, Robert L.
; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 1871-129
; CURRENT APPLICATION NUMBER: US/09/509,196A
; CURRENT FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: P09388
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: PCT AU98/00795
; PRIOR FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1074
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-196A-2

Query Match 86.7%; Score 5602; DB 9; Length 1074;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1068; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy        167 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 60
```

```
Qy        227 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 120

Qy        287 TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNA 346
              |||||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db        121 TPLHLAAGYNRVKIVQLLLQHGRDVHAKDKGDLVPLHNACSYGHYEVTELLVKHGGCVNA 180

Qy        347 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKG 406
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        181 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCKNKSAIDLAPTPQLKERLAYEFKG 240

Qy        407 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 300

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 526
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 360

Qy        527 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 586
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 420

Qy        587 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480

Qy        647 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 540

Qy        707 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 600

Qy        767 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEA 826
              |||||||||||||||||||||||||||||||||||||||||| :||||||||||||||||
Db        601 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFTPLHEA 660

Qy        827 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 720

Qy        887 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 780

Qy        947 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 840

Qy       1007 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 900

Qy       1067 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 960

Qy       1127 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1020

Qy       1187 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||||||:||||||||||||||||||||||
Db       1021 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEEAYPEYLITYQIMRPEGMVDG 1074

RESULT 10
US-09-972-115A-4
; Sequence 4, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 1267
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: The 'Xaa' at location 1 stands for Thr, Ala, Pro, or Ser.
;   NAME/KEY: misc_feature
;   LOCATION: (42)..(42)
;   OTHER INFORMATION: The 'Xaa' at location 42 stands for Lys, Arg, Thr, or Met.
;   NAME/KEY: misc_feature
;   LOCATION: (98)..(98)
;   OTHER INFORMATION: The 'Xaa' at location 98 stands for Lys, Arg, Thr, or Met.
;   NAME/KEY: misc_feature
;   LOCATION: (110)..(110)
;   OTHER INFORMATION: The 'Xaa' at location 110 stands for Thr, Ala, Pro, or Ser.
;   NAME/KEY: misc_feature
;   LOCATION: (114)..(114)
;   OTHER INFORMATION: The 'Xaa' at location 114 stands for Asn, Ser, Thr, or Ile.
;   NAME/KEY: misc_feature
;   LOCATION: (126)..(126)
;   OTHER INFORMATION: The 'Xaa' at location 126 stands for Lys, or Asn.
;   NAME/KEY: misc_feature
;   LOCATION: (139)..(139)
;   OTHER INFORMATION: The 'Xaa' at location 139 stands for Ile, Val, Leu, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (144)..(144)
;   OTHER INFORMATION: The 'Xaa' at location 144 stands for Thr.
;   NAME/KEY: misc_feature
;   LOCATION: (153)..(153)
;   OTHER INFORMATION: The 'Xaa' at location 153 stands for Thr.
;   NAME/KEY: misc_feature
;   LOCATION: (166)..(166)
;   OTHER INFORMATION: The 'Xaa' at location 166 stands for Ser, Gly, Arg, or Cys.
;   NAME/KEY: misc_feature
;   LOCATION: (173)..(173)
;   OTHER INFORMATION: The 'Xaa' at location 173 stands for Ile, Val, Leu, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (174)..(174)
;   OTHER INFORMATION: The 'Xaa' at location 174 stands for Glu, or Lys.
;   NAME/KEY: misc_feature
;   LOCATION: (183)..(183)
;   OTHER INFORMATION: The 'Xaa' at location 183 stands for Val.
;   NAME/KEY: misc_feature
;   LOCATION: (185)..(185)
;   OTHER INFORMATION: The 'Xaa' at location 185 stands for Thr, Ala, Pro, or Ser.
;   NAME/KEY: misc_feature
;   LOCATION: (188)..(188)
;   OTHER INFORMATION: The 'Xaa' at location 188 stands for Asn, Asp, His, or Tyr.
;   NAME/KEY: misc_feature
;   LOCATION: (202)..(202)
;   OTHER INFORMATION: The 'Xaa' at location 202 stands for His, or Pro.
;   NAME/KEY: misc_feature
;   LOCATION: (204)..(204)
;   OTHER INFORMATION: The 'Xaa' at location 204 stands for Glu, or Lys.
;   NAME/KEY: misc_feature
```

; LOCATION: (205)..(205)
; OTHER INFORMATION: The 'Xaa' at location 205 stands for Asn, Ser, Thr, or Ile.
; NAME/KEY: misc_feature
; LOCATION: (207)..(207)
; OTHER INFORMATION: The 'Xaa' at location 207 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (211)..(211)
; OTHER INFORMATION: The 'Xaa' at location 211 stands for Gln, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (213)..(213)
; OTHER INFORMATION: The 'Xaa' at location 213 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (215)..(215)
; OTHER INFORMATION: The 'Xaa' at location 215 stands for Lys, Asn, Glu, Asp, Gln, His
; OTHER INFORMATION: , a stop codon, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (227)..(227)
; OTHER INFORMATION: The 'Xaa' at location 227 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (228)..(228)
; OTHER INFORMATION: The 'Xaa' at location 228 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (237)..(237)
; OTHER INFORMATION: The 'Xaa' at location 237 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (238)..(238)
; OTHER INFORMATION: The 'Xaa' at location 238 stands for Asn, Ser, Thr, Ile, Asp, Gly
; OTHER INFORMATION: , Ala, Val, His, Arg, Pro, Leu, Tyr, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (667)..(667)
; OTHER INFORMATION: The 'Xaa' at location 667 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (725)..(725)
; OTHER INFORMATION: The 'Xaa' at location 725 stands for Lys, or Asn.
; NAME/KEY: misc_feature
; LOCATION: (727)..(727)
; OTHER INFORMATION: The 'Xaa' at location 727 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (735)..(735)
; OTHER INFORMATION: The 'Xaa' at location 735 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (741)..(741)
; OTHER INFORMATION: The 'Xaa' at location 741 stands for Met, Val, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (763)..(763)
; OTHER INFORMATION: The 'Xaa' at location 763 stands for Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (764)..(764)
; OTHER INFORMATION: The 'Xaa' at location 764 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (787)..(787)
; OTHER INFORMATION: The 'Xaa' at location 787 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (788)..(788)
; OTHER INFORMATION: The 'Xaa' at location 788 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (789)..(789)
; OTHER INFORMATION: The 'Xaa' at location 789 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (790)..(790)
; OTHER INFORMATION: The 'Xaa' at location 790 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (791)..(791)
; OTHER INFORMATION: The 'Xaa' at location 791 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (792)..(792)
; OTHER INFORMATION: The 'Xaa' at location 792 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (793)..(793)
; OTHER INFORMATION: The 'Xaa' at location 793 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (794)..(794)
; OTHER INFORMATION: The 'Xaa' at location 794 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (795)..(795)
; OTHER INFORMATION: The 'Xaa' at location 795 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (796)..(796)
; OTHER INFORMATION: The 'Xaa' at location 796 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: The 'Xaa' at location 797 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (798)..(798)
; OTHER INFORMATION: The 'Xaa' at location 798 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (799)..(799)
; OTHER INFORMATION: The 'Xaa' at location 799 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (800)..(800)
; OTHER INFORMATION: The 'Xaa' at location 800 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (801)..(801)
; OTHER INFORMATION: The 'Xaa' at location 801 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (802)..(802)

Query Match 84.3%; Score 5452; DB 10; Length 1267;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 1072; Conservative 6; Mismatches 161; Indels 4; Gaps 2;

```
Qy          1 RCSARRGAAGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 60
              ||  |||||||||| |     | ||||||||||||||||||||||||||||||||||||
Db         26 RCLRRRGAAGGQGAHRXGARGRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLL 85

Qy         61 RLLALLLAVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTP 120
              |||||||||||| |||||||||| ||| |||||||||| |||||||||||| |:|| |
Db         86 RLLALLLAVAAAXIMSGRRCAGGGXACAXAAAEAVEPAARXLFEACRNGDVERXKKLVXP 145

Qy        121 EKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEV 180
              ||||||| |||||||||| | |||||:  ||| |||| | || ||||||||||||| |
Db        146 EKVNSRDXAGRKSTPLHFPAXFGRKDLXXYLLTNGANXQXRDXGGLIPLHNACSFGXAXX 205

Qy        181 VNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP 240
              : ||| | | |||||||||||  ||||||||  ||||||||||||||||||||||||||||
Db        206 IXLLLXHXAXPNARDNWNYTPXXEAAIKGKIXXCIVLLQHGAEPTIRNTDGRTALDLADP 265
```

```
Qy        241 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 SAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKI 325

Qy        301 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 VQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAAS 385

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 445

Qy        421 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 505

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 565

Qy        541 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 625

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        626 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVN 685

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 720
              |||||||||||||||||||||||||||||||| : |   ||||||| ||||| ||||
Db        686 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKKKKGXIXLDLVKDGXTDIQDXLRGD 745

Qy        721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 780
              | ||||||||||||||   |||||||||||||||||| ||
Db        746 AVLLDAAKKGCLARVKKXXFPDNVNCRDTQGRHSTPLHL-AGXXXXXXXXXXXXXXXXXX 804

Qy        781 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 864

Qy        841 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 900
                                      ||||||| |||| ||||||||||||||||||||
Db        865 XXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALS 924

Qy        901 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        925 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 984

Qy        961 LEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTL 1020
              |||||||| |||||||||||||||||||| |||||||||| |||||||||||||||||||
Db        985 LEHLMDIFXREQITLDVLVEMGHKELKEIXINAYGHRHKLIKSFERLISGQQGLNPYLTL 1044

Qy       1021 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK--- 1077
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db       1045 NTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNRAKI 1104

Qy       1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
              :  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1105 RHEERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1164

Qy       1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1224

Qy       1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              |||||||||||||||||||||||||||||||||||||||||||
Db       1225 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1267


RESULT 11
US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-841-835-2

  Query Match           79.0%;  Score 5103.5;  DB 9;  Length 1327;
  Best Local Similarity  79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|    :  : ||      : |     |  :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||       ||  | ||| |||||||| ||||||    ||::| |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||||:|||  |||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||||:|||||||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||:|||||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              ||||||||||||||:|||||||||||||||||||||||||||||||||||:||||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
```

```
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501

Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561

Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621

Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681

Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741

Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801

Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861

Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921

Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979

Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039

Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099

Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159

Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219

Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235

Db       1299 QAYPEYLITYQIMKPE 1314


RESULT 12
US-09-972-115A-8
; Sequence 8, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 1327
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-8

  Query Match           79.0%;  Score 5103.5;  DB 10;  Length 1327;
  Best Local Similarity  79.2%;  Pred. No. 0;
  Matches  963;  Conservative  110;  Mismatches  128;  Indels   15;  Gaps    3;

Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81

Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141

Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201

Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261

Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321

Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381

Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441

Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501

Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561

Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621

Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681

Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741

Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801

Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861

Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
```

```
Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979

Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039

Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099

Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159

Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219

Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235

Db       1299 QAYPEYLITYQIMKPE 1314
```

RESULT 13
US-10-199-937-4
; Sequence 4, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-4

Query Match 79.0%; Score 5103.5; DB 14; Length 1327;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 963; Conservative 110; Mismatches 128; Indels 15; Gaps 3;

```
Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81

Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141

Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201

Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261

Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321

Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381

Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441

Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501

Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561

Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621

Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681

Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741

Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 801

Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDI 885

Qy        802 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861

Db        886 AALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATA 945

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921

Db        946 DDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTGP 998

Qy        922 FSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 979

Db        999 LAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVLA 1058

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEF 1039

Db       1059 DMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEY 1118

Qy       1040 QSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHAN 1099

Db       1119 QSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHN 1178

Qy       1100 ERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK 1159

Db       1179 ERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHK 1238

Qy       1160 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1219

Db       1239 DRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGE 1298

Qy       1220 QAYPEYLITYQIMRPE 1235

Db       1299 QAYPEYLITYQIMKPE 1314
```

RESULT 14
US-09-972-115A-2
; Sequence 2, Application US/09972115A
; Publication No. US20030032769A1

```
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1333
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: The 'Xaa' at location 1 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: The 'Xaa' at location 2 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: The 'Xaa' at location 3 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: The 'Xaa' at location 4 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: The 'Xaa' at location 5 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (6)..(6)
;   OTHER INFORMATION: The 'Xaa' at location 6 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: The 'Xaa' at location 7 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: The 'Xaa' at location 8 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: The 'Xaa' at location 9 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: The 'Xaa' at location 10 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (11)..(11)
;   OTHER INFORMATION: The 'Xaa' at location 11 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (12)..(12)
;   OTHER INFORMATION: The 'Xaa' at location 12 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: The 'Xaa' at location 13 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (14)..(14)
;   OTHER INFORMATION: The 'Xaa' at location 14 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (15)..(15)
;   OTHER INFORMATION: The 'Xaa' at location 15 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (16)..(16)
;   OTHER INFORMATION: The 'Xaa' at location 16 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (17)..(17)
;   OTHER INFORMATION: The 'Xaa' at location 17 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (19)..(19)
;   OTHER INFORMATION: The 'Xaa' at location 19 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (20)..(20)
;   OTHER INFORMATION: The 'Xaa' at location 20 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (21)..(21)
;   OTHER INFORMATION: The 'Xaa' at location 21 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (22)..(22)
;   OTHER INFORMATION: The 'Xaa' at location 22 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (23)..(23)
;   OTHER INFORMATION: The 'Xaa' at location 23 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: The 'Xaa' at location 24 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (25)..(25)
;   OTHER INFORMATION: The 'Xaa' at location 25 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (26)..(26)
;   OTHER INFORMATION: The 'Xaa' at location 26 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (27)..(27)
;   OTHER INFORMATION: The 'Xaa' at location 27 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (28)..(28)
;   OTHER INFORMATION: The 'Xaa' at location 28 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (29)..(29)
;   OTHER INFORMATION: The 'Xaa' at location 29 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (30)..(30)
;   OTHER INFORMATION: The 'Xaa' at location 30 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (31)..(31)
;   OTHER INFORMATION: The 'Xaa' at location 31 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (32)..(32)
;   OTHER INFORMATION: The 'Xaa' at location 32 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (33)..(33)
;   OTHER INFORMATION: The 'Xaa' at location 33 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (34)..(34)
;   OTHER INFORMATION: The 'Xaa' at location 34 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (35)..(35)
;   OTHER INFORMATION: The 'Xaa' at location 35 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (36)..(36)
;   OTHER INFORMATION: The 'Xaa' at location 36 stands for Lys, Asn, Arg, Ser, Thr, Ile,

  Query Match           77.8%;  Score 5031.5;  DB 10;  Length 1333;
  Best Local Similarity  86.7%;  Pred. No. 0;
  Matches  983;  Conservative    7;  Mismatches  135;  Indels    9;  Gaps    2;

Qy       116 RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSF 175
            ||||||||  :| |||||||| ||||||:  :|||||||  |   :||||||| ||||
Db       200 RLVTPEKVNXXNTXGRKSTPLHFPAGFGRKNXXKYLLQNGANXQXLYNGGLIPLHXACSF 259

Qy       176 GHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTAL 235
             |||: :|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       260 GHAKXINLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTAL 319

Qy       236 DLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY 295
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       320 DLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY 379

Qy       296 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 355
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       380 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 439

Qy       356 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       440 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 499

Qy       416 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       500 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 559

Qy       476 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       560 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 619

Qy       536 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       620 ISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 679

Qy       596 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 655
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       680 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKH 739

Qy       656 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD 715
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| |||||
Db       740 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTLLDLVKDGXTDIQD 799

Qy       716 LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA--AGYNNLEVAEYLL 773
             |||||  |||||||||||||||   |||||||||||||||||||
Db       800 XLRGDAVLLDAAKKGCLARVKKXXFPDNVNCRDTQGRHSTPLHLAGXXXXXXXXXXXXXX 859

Qy       774 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 833

Db       860 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 919

Qy       834 LCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPG 893
                                                  ||||||| |||| ||||||||||||
Db       920 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPG 979

Qy       894 ATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSIT 953
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       980 ATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSIT 1039

Qy       954 QFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 1013
             |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||||
Db      1040 QFVRNLGLEHLMDIFXREQITLDVLVEMGHKELKEIXINAYGHRHKLIKGVERLISGQQG 1099

Qy      1014 LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQK 1073
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1100 LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQK 1159

Qy      1074 VCNK-------KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1126
             |||         :  ||||||||||||||||||||||||||||||||||||||||||||||
Db      1160 VCNXXXXXAKIRHEERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 1219

Qy      1127 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1220 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1279

Qy      1187 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1280 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1333

RESULT 15
US-10-199-937-89
; Sequence 89, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
```

```
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
;   LENGTH: 784
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-89

  Query Match             63.5%;  Score 4102;  DB 14;  Length 784;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  784;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        457 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 60

Qy        517 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 120

Qy        577 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 180

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 240

Qy        697 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 300

Qy        757 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 360

Qy        817 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 876
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 420

Qy        877 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 936
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 480

Qy        937 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 996
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 540

Qy        997 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 600

Qy       1057 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 660

Qy       1117 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 720

Qy       1177 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 780

Qy       1237 MVDG 1240
              ||||
Db        781 MVDG 784
```

```
RESULT 16
US-10-199-937-139
; Sequence 139, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 139
;   LENGTH: 1181
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-10-199-937-139

  Query Match             61.8%;  Score 3997;  DB 14;  Length 1181;
  Best Local Similarity   65.2%;  Pred. No. 0;
  Matches  771;  Conservative  166;  Mismatches  190;  Indels   56;  Gaps   13;

Qy         91 AAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEY 150
              |   | :|  ||| |||: |:: :||:|:||: ||:|||||||||||||||||||:||::|||:
Db         17 AVMANDP-LRELSEACKTGEIAKVKKLITPQTVNARDTAGRKSTPLHFAAGYGRREVVEF 75

Qy        151 LLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGK 210
              || :||::|| |:||| |||| ||||||||| |||: || ||  |||||||||||| |||
Db         76 LLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYTPLHEAASKGK 135

Qy        211 IDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMAL 270
              :|||: ||||||  ||||:: :| |:||| : : ||||||:||||||:|||| |::::||
Db        136 VDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAARSGAEDRLLAL 195

Qy        271 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGH 330
              |||||||||||||:|||||||||||: ||::|| :|||||||||| ||||||||||
Db        196 LTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLVPLHNACSYGH 255

Qy        331 YEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL 390
              ::||:|::: || ||| ||| |||||||||||:|||||||||| ||||||||||:|||||
Db        256 FDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLLNCHSKSAIDA 315

Qy        391 APTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPY 450
              ||| :|:||:|:|:||| || | |: ||:| || :  |:||| || | :| || |  ||
Db        316 APTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDTPLHLAVVSPD 375

Qy        451 PKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTS 510
              ||||: ||| |||: :||| | ||||||:|:|  | | :||::|  |||||||:||||
Db        376 GKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKVNALDSLGQTP 435

Qy        511 LHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQL 570
              ||| |     | |||||  | ||:||:| || |: :::| :||:       :::  |
Db        436 LHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK-----NPPDSETHL 488

Qy        571 LEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 628
              ||||||||::||:::      ||||||::|| |||||||||:||| ||::||:|||:|:|
Db        489 LEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQFLLEHGAEVYA 548

Qy        629 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 688
              |||||||||||||||||| |||||||| |||:|||||||||||||||:||||||:|
Db        549 ADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGKYDICKLLLKH 608

Qy        689 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 748
              |||| |||||| || |||||: | |: :|||| :|||||||| ||||::| :|:::|||||
Db        609 GADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRLVTPESINCRD 668
```

```
Qy        749 TQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKY 808
Db        669 AQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGHLDIAALLIKH 728

Qy        809 NACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL 868
Db        729 KTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIELATADDVKCLL 788

Qy        869 TAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAASSLDNLSGSF 922
Db        789 QDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPTTETVLLPTGA 847

Qy        923 SELSSV-----VSSS------------GTEGASSLEKKEVPGVD--FSITQFVRNLGLEH 963
Db        848 SMILSVPVPLPLSSSTRISPAQGAEANGAEGSSSDDL--LPDADTITNVSGFLSSQQLHH 905

Qy        964 LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTS 1023
Db        906 LIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS------------TT 953

Qy       1024 GSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK 1077
Db        954 GIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYNIIRVQKVQNR 1013

Qy       1078 KLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAE 1137
Db       1014 KLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGGMFGAGIYFAE 1073

Qy       1138 NSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHH 1197
Db       1074 HSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAMKMAHAPPGHH 1133

Qy       1198 SVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db       1134 SVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176
```

RESULT 17
US-10-199-937-91
; Sequence 91, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-91

Query Match 61.2%; Score 3959; DB 14; Length 756;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        485 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 544
Db          1 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 60

Qy        545 QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPL 604
```

```
Db         61 QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPL 120

Qy        605 HFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL 664
Db        121 HFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL 180

Qy        665 WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALL 724
Db        181 WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALL 240

Qy        725 DAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 784
Db        241 DAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 300

Qy        785 GGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGAD 844
Db        301 GGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGAD 360

Qy        845 PTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPS 904
Db        361 PTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPS 420

Qy        905 SPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 964
Db        421 SPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 480

Qy        965 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 1024
Db        481 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 540

Qy       1025 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 1084
Db        541 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 600

Qy       1085 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 1144
Db        601 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 660

Qy       1145 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 1204
Db        661 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 720

Qy       1205 VNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
Db        721 VNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 756
```

RESULT 18
US-09-964-899-41
; Sequence 41, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-41

Query Match 60.9%; Score 3938; DB 9; Length 802;
Best Local Similarity 67.2%; Pred. No. 2.4e-317;
Matches 787; Conservative 4; Mismatches 5; Indels 376; Gaps 5;

```
Qy         75 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST 60

Qy        135 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR 120

Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGE----- 249
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        121 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGKSVYSG 180

Qy        250 YKKDELLES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHG 308
              | :  :  : ||||||||||||||||||||||||
Db        181 YSRKPVKNNLARSGNEEKMMALLTPLNVNCHASDGR----------------------- 216

Qy        309 ADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCS 368
                                                |||||||||||||||||||||||||||
Db        217 -----------------------------KHGACVNAMDLWQFTPLHEAASKNRVEVCS 246

Qy        369 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE 306

Qy        429 MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHND 488
              |||||||||||||
Db        307 MVNFKHPQTHETAL---------------------------------------------- 320

Qy        489 VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN 548

Db        321 ------------------------------------------------------------ 320

Qy        549 ENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAA 608
                                                |||||||||||||||||||||||||
Db        321 ----------------------------------KLCTVQSVNCRDIEGRQSTPLHFAA 345

Qy        609 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 405

Qy        669 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 465

Qy        729 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 525

Qy        789 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 848
              |||||||||
Db        526 PLHNAASYG---------------------------------------------------- 534

Qy        849 NQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSS 908

Db        535 ------------------------------------------------------------ 534

Qy        909 LSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIF 968

Db        535 ------------------------------------------------------------ 534

Qy        969 EREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTI 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ----ITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTI 590

Qy       1029 LIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        591 LIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 650

Qy       1089 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        651 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 710

Qy       1149 IGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGL 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 IGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGL 770

Qy       1209 ALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||
Db        771 ALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 802
```

RESULT 19
US-09-841-835-10
; Sequence 10, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 949 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-841-835-10

Query Match 51.2%; Score 3307; DB 9; Length 949;
Best Local Similarity 76.2%; Pred. No. 8e-265;
Matches 638; Conservative 72; Mismatches 119; Indels 8; Gaps 2;

```
Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
              :| | ||:|  :||  : : |:|     :   : ||      : |     | :|
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
              | |||      ||  | ||| |||||||| |||||    ||::| |||||:|||||||
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
              ||||||||:||| |||| |||||||||||||||||||||:||| |||||||||||||||
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
              |||||||||||||||||||:| |||||||::|||||||||||||||||||||||:|||
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345
```

```
Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
              |||||:|||||||||||||||||||||||||||||||||:||||||||||||||||||| |||
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||:|||||:|
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
              || |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:||||
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
              |||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:|
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
              ||| ||||:|||||  |||||||||||| ||:||||||| ||||| |||:| |   :
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPI 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 621
              |: | :||||:||||:||||:|| |:|||||:||| |||||||||||||||||||
Db        646 RTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLH 705

Qy        622 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 681
              ||||||||||||||||||||||||||||:||| |||||||||||||||||||||||||||
Db        706 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEI 765

Qy        682 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 741
              |||||:|||||||||||||||||||:|||||||||:|||||||||||||||||:|| :|
Db        766 CKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCTP 825

Qy        742 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH--V 799
              :|:|||||||:|||||||||||||||||||:|||||||||||||||||||||||||
Db        826 ENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCLA 885

Qy        800 DVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856
              | |  |   |  |  |||| ||    ::  || ||||  ::: |  ||
Db        886 RVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942
```

RESULT 20
US-10-199-937-99
; Sequence 99, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-99

Query Match 42.1%; Score 2722; DB 14; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.1e-216;
Matches 522; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         72 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR 60

Qy        132 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADP 120

Qy        192 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 NARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYK 180

Qy        252 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        181 KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV 240

Qy        312 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 HAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLL 300

Qy        372 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        301 SYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN 360

Qy        432 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVE 420

Qy        492 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        421 VVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 480

Qy        552 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 593
              |||||||||||||||||||||||||||||||||||||||||||
Db        481 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 522
```

RESULT 21
US-10-276-774-1690
; Sequence 1690, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1690
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-1690

Query Match 33.5%; Score 2165; DB 12; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.3e-170;
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        348 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 407
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 60

Qy        408 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 467
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 120

Qy        468 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 527
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 180
```

```
Qy        528 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 587
Db        181 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 240

Qy        588 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 647
Db        241 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 300

Qy        648 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
Db        301 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 360

Qy        708 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 760
Db        361 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 413
```

RESULT 22
US-09-841-835-8
; Sequence 8, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 673 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-841-835-8

Query Match 33.0%; Score 2132.5; DB 9; Length 673;
Best Local Similarity 73.3%; Pred. No. 1.4e-167;
Matches 417; Conservative 54; Mismatches 91; Indels 7; Gaps 2;

```
Qy         22 AAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRCA 81
Db        112 SAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPG------SSLAESPEAAGVSSTAPL 165

Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAG 141
Db        166 GPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG 225

Qy        142 FGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 201
Db        226 FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTP 285

Qy        202 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 261
Db        286 LHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARS 345

Qy        262 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 321
Db        346 GNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVP 405

Qy        322 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 381
Db        406 LHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVN 465

Qy        382 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 441
Db        466 CHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETA 525

Qy        442 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 501
Db        526 LHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMN 585

Qy        502 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 561
Db        586 ALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYGS 645

Qy        562 GNSEADRQLLEAAKAGDVETVKKLCTVQS 590
Db        646 DPSIISLQGFTAAQMGN-EAVQQILSGHS 673
```

RESULT 23
US-10-199-937-178
; Sequence 178, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
; LENGTH: 1099
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Parpla-Tank2b
; OTHER INFORMATION: Fusion
US-10-199-937-178

Query Match 31.8%; Score 2056.5; DB 14; Length 1099;
Best Local Similarity 62.1%; Pred. No. 6.6e-161;
Matches 459; Conservative 38; Mismatches 115; Indels 127; Gaps 17;

```
Qy        544 LQMGNENVQQLLQEGISLG---NSEADRQLLEAAKAGDVETV-KKLCTVQSVNCRDIEGR 599
Db        437 LSRNKDEVKAMIEK---LGGKLTGTANKASLCISTKKEVEKMNKKMEEVKEANIR----- 488

Qy        600 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 659
```

```
Db        489 --------------VVSEDFLQ---DVSASTK----------SLQELFLAHILSPWGAEV 521

Qy        660 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 719
                  |  |:   | :||          ||  :||:: |    : :   :  ::  |:|
Db        522 ------KAEPVEVVAPRGK---------SGAALSKKSK-GQVKEEGINKSEKRMKLTLKG 565

Qy        720 DAAL-----------------LDAAKKGCLARVK--------KLSSPDNVN---CRDTQ 750
                ||:                 : :|  | :  ||        :|   |  |      :
Db        566 GAAVDPDSGLEHSAHVLEKGGKVPSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSW 625

Qy        751 GRHSTPLHLAAGYNNLE---VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 807
                ||  | :    | | ||      |  ::|   :  :  |      |   |
Db        626 GRVGTVI----GSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSKNFTKYPK--------- 672

Qy        808 YNACVNATDKWAFTPLH------EAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 861
                          | ||      | | |  |          :|  |  ||||||||||
Db        673 -----------KFYPLEIDYGQDEEAVKKLT----------VNPGTKRVEGQTPLDLVSA 711

Qy        862 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 921
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 771

Qy        922 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 981
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 831

Qy        982 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1041
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 891

Qy       1042 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 1101
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 951

Qy       1102 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1161
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1011

Qy       1162 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1221
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1071

Qy       1222 YPEYLITYQIMRPEGMVDG 1240
                ||||||||||||||||||
Db       1072 YPEYLITYQIMRPEGMVDG 1090
```

RESULT 24
US-10-199-937-160
; Sequence 160, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-160

Query Match 29.3%; Score 1895; DB 14; Length 362;
Best Local Similarity 100.0%; Pred. No. 2.6e-148;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        352 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 411
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 60

Qy        412 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 471
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 120

Qy        472 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 531
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 180

Qy        532 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 591
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 240

Qy        592 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 651
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 300

Qy        652 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 711
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 360

Qy        712 DI 713
                ||
Db        361 DI 362
```

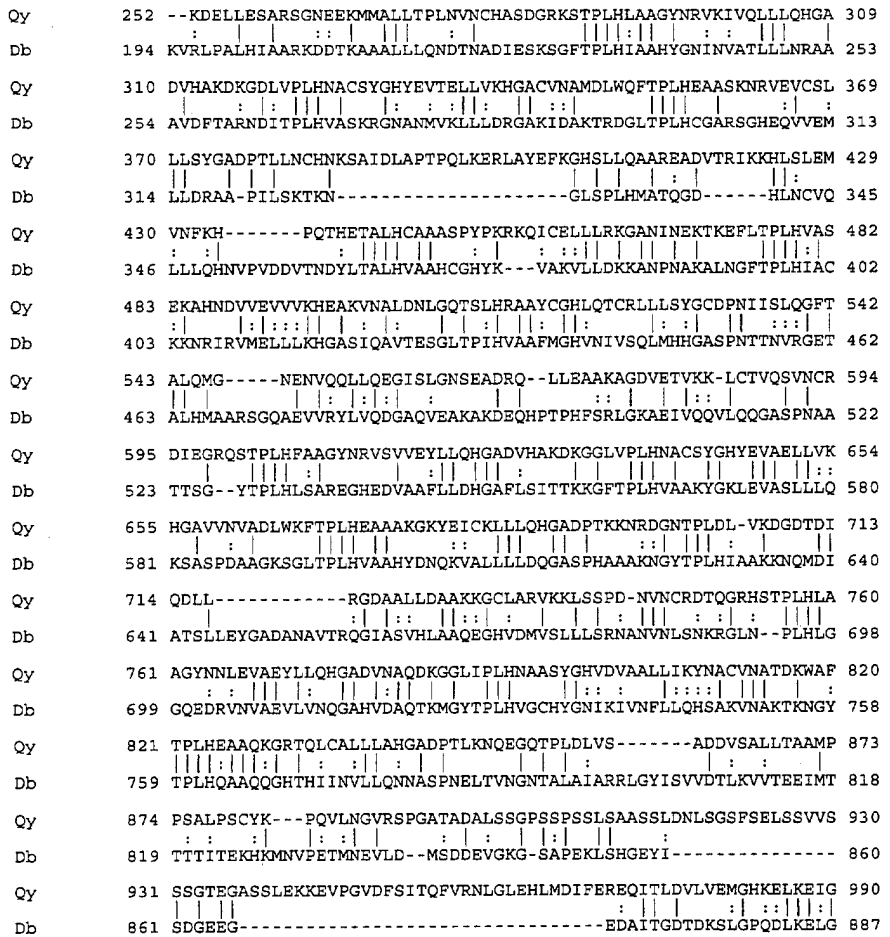
RESULT 25
US-10-205-194-117
; Sequence 117, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; CURRENT FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 1762
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Ankyrin isoform
US-10-205-194-117

Query Match 13.7%; Score 888.5; DB 14; Length 1762;
Best Local Similarity 30.0%; Pred. No. 1.3e-63;
Matches 288; Conservative 133; Mismatches 396; Indels 143; Gaps 25;

```
Qy         82 GGGAACASAAAEAVEPAARELFEACRNGDVER----VKRLVTPEKVNSRDTAGRKSTPLH 137
                |   :| | :::  :  |  |  | | | :|:    :|  |     ||  :  |  :  ||
Db         20 GDFGSCYSRSSQKSDANASYL-RAARAGHLEKALDYIKNGV---DVNICNQNGLNA--LH 73

Qy        138 FAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNW 197
                |:  |  :||  |||  |||       |   || |   | |||| :|: :||: ||:
Db         74 LASKEGHVEVVSELLQREANVDQPTKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQN 133

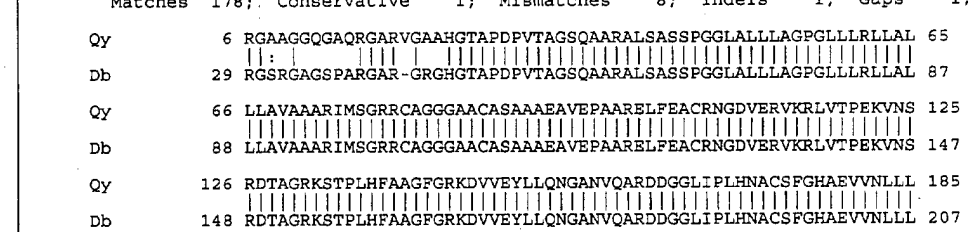
Qy        198 NYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK- 251
                 :|||: || :  ::|   || :||  ::   || | | :|      |     :|  : |
Db        134 GFTPLYMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKG 193
```



RESULT 26
US-10-199-937-105
; Sequence 105, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0

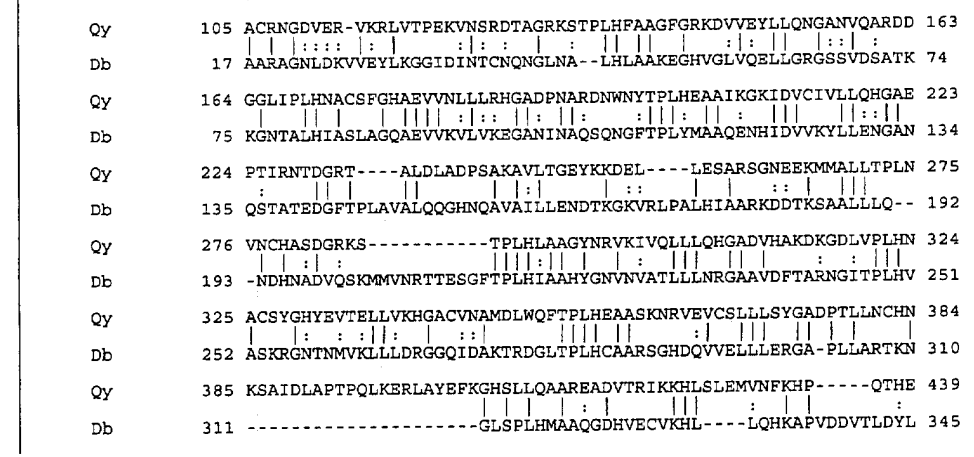
; SEQ ID NO 105
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-105

Query Match 13.7%; Score 886.5; DB 14; Length 215;
Best Local Similarity 94.7%; Pred. No. 5.6e-65;
Matches 178; Conservative 1; Mismatches 8; Indels 1; Gaps 1;



RESULT 27
US-10-108-260A-3237
; Sequence 3237, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3237
; LENGTH: 1330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3237

Query Match 13.6%; Score 877; DB 15; Length 1330;
Best Local Similarity 29.7%; Pred. No. 7.5e-63;
Matches 293; Conservative 137; Mismatches 402; Indels 156; Gaps 23;



```
Qy        440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
              |||| ||   :    :: :||| | || | :     ||||:| :|    :|::||: |
Db        346 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAS 402

Qy        500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN-----ENVQQL 554
              : |:   | | : | ||: |||    |||  |  |:: :::| ||| |       | |: |
Db        403 IQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCL 462

Qy        555 LQEGI----------------SLGNSEADRQLLE----------------AAKAGDV 579
              |: |                || :|  : ||:                :|: | |
Db        463 LRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTNRYTPLHISAREGQV 522

Qy        580 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
              :    |    : :    : :  |||| || |  : | : |||  |   :  | || |||
Db        523 DVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHV 581

Qy        640 ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDG 699
              |  | : :|| ||:: ||  :      :|||| || | : :|   || :||:     : |
Db        582 AAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNIVTKQG 641

Qy        700 NTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLH 758
              |||  |  ::|  ||:  ||   | :  : | |                       | ||
Db        642 VTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----------------------TSLH 679

Qy        759 LAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
              |||  : : ||: | :|||| :|  | |  ||  |  ||:| :   |:|  | |||  |
Db        680 LAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYGNVKMVNFLLKQGANVNAKTKN 739

Qy        819 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLTAAM 872
              :|||||:|||:| |  :  :|| ||| |      | | | :      :|  |   ::|  :
Db        740 GYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALAIAKRLGYISVVDTLKVVTEEV 799

Qy        873 PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSEL------- 925
              :       |    : :  |    : |         :|      | ::|   |
Db        800 TTTTTTITEK----HKLNVPETMTEVL--------DVSDEEGDDTMTGDGGEYLRPEDLK 847

Qy        926 ----SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RNLGLEHLMDIFEREQITLDVLV 979
                  |: ||   :| : |      |   |:  |   |:  | |    : |:   :|  |
Db        848 ELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRSHTLSHAS--YLRDSAVMDDSV 905

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERL 1007
              :   ::  :   |  : ::|  | | |
Db        906 VIPSHQVSTLAKEAERNSYRLSWGTENL 933
```

RESULT 28
US-09-972-115A-20
; Sequence 20, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-20

Query Match 13.5%; Score 874; DB 10; Length 160;
Best Local Similarity 100.0%; Pred. No. 3.7e-64;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 60

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 120

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||
Db        121 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 160
```

RESULT 29
US-10-199-937-46
; Sequence 46, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence not
; OTHER INFORMATION: specified as protein-coding is vector sequence
US-10-199-937-46

Query Match 13.5%; Score 874; DB 14; Length 160;
Best Local Similarity 100.0%; Pred. No. 3.7e-64;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1081 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 60

Qy       1141 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 120

Qy       1201 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1240
              ||||||||||||||||||||||||||||||||||||||||
Db        121 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 160
```

RESULT 30
US-09-964-899-43
; Sequence 43, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309

```
;  PRIOR FILING DATE: 2001-06-14
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
;   LENGTH: 1724
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-09-964-899-43
```

Query Match 12.6%; Score 815.5; DB 9; Length 1724;
Best Local Similarity 25.3%; Pred. No. 1.5e-57;
Matches 286; Conservative 147; Mismatches 380; Indels 319; Gaps 27;

```
Qy        136 LHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD 195
Db         40 LHLASKEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQS 99

Qy        196 NWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDEL 255
Db        100 QNGFTPLYMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVA---------------- 142

Qy        256 LESARSGNEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAK 314
Db        143 ---LQQGHDQVVSLLLE------NDTKGKVRLPALHIAARKDDTKAAALLLQNDNNADVE 193

Qy        315 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 374
Db        194 SKSGFTPLHIAAHYGNINVATLLLNRAAAVDFTARNDITPLHVASKRGNANMVKLLLDRG 253

Qy        375 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
Db        254 A-----------KIDAKTRDGL---------------------------------------- 264

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
Db        265 -----TPLHCGARSGH---EQVVEMLLDRAAPILSKTKNGLSPLHMATQGDHLNCVQLLL 316

Qy        495 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 554
Db        317 QHNVPVDDVTNDYLTALHVAAHCGHYKVAKVLLDKKANPNAKALNGF---------------- 363

Qy        555 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 614
Db        364 ------------------------------------------------TPLHIACKKNRIK 376

Qy        615 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 674
Db        377 VMELLLKHGASIQAVTERGETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISA 436

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLA 733
Db        437 RLGKADIVQQLLQQGASPNAATTSGYTPLHLSAREGHEDVAAFLLDHGASLSITTKSGL- 495

Qy        734 RVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNA 793
Db        496 --------------------TPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIA 534

Qy        794 ASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQ 853
Db        535 AKKNQMDIATTLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNANVNLSNKSGL 594

Qy        854 TPLDLVSADD----VSALLTAAMPPSALPSCYKPQVLNG------VRSPGAT-------- 895
Db        595 TPLHLAAQEDRVNVAEVLVNQGAHVDAQTKVYGPPLPHGKECVHLVTANGYTPLHQAAQQ 654

Qy        896 -----ADALSSGPSSPSSLSAASS-------------LDNLSGSFSELSSVVSSSGTEG 936
Db        655 GHTHIINVLLQNNASPNELTVTVTEKHKMNVPETMNEVLDMSDDEVRKANAPEMLSDGEY 714

Qy        937 ASSLEK------KEVPGV-DFSITQFVRNLGLEHLMDIFEREQITLDVLVE--------- 980
Db        715 ISDVEEGNRCTWYKIPKVQEFTVK----------TDTFKREAFDVGLLSTSAGEDAMTG 763

Qy        981 -----MGHKELKEIGINAY----------------------------GHRHKLIKGV 1004
Db        764 DTDKYLGPQDLKELGDDSLPAEGYMGFSLGARSARFLVSFMVDARGGSMRGSRHHGMRII 823

Qy       1005 E-------------RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS-VEEEMQSTV 1050
Db        824 IPPRKCTAPTRITCRLVKRHKLANPPPMVEGEGLASRLVEMGPAGAQFLGPVIVEIP--- 880

Qy       1051 REHRDGGHAGGIFNRYNILKIQKVCNKKLWE--RYTHRRKEVSE-----ENHNHANERML 1103
Db        881 -------HFGSMRGKERELIVLRSENGETWKEHQFDSKNEDLTELLNGMDEELDSPEEL- 932

Qy       1104 FHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENS---SKSNQYVYGIGGG 1152
Db        933 --GKKRICRIITKDFPQ-------------YFAVVSRIKQESNQ--IGPEGG 967
```

RESULT 31
US-10-106-698-6500
; Sequence 6500, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6500
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6500

Query Match 11.9%; Score 772; DB 14; Length 153;
Best Local Similarity 99.3%; Pred. No. 1e-55;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        924 ELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGH 983
Db          1 ELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVLVEMGH 60

Qy        984 KELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE 1043
Db         61 KELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE 120

Qy       1044 EEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 1076
Db        121 EEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 153
```

RESULT 32
US-10-221-278-343
; Sequence 343, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267

```
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 343
;   LENGTH: 1053
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-221-278-343

  Query Match           11.8%;  Score 760.5;  DB 12;  Length 1053;
  Best Local Similarity  27.6%;  Pred. No. 2.4e-53;
  Matches  275;  Conservative  129;  Mismatches  357;  Indels  235;  Gaps   33;

Qy        102 LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
              | :|  ||| : |: |:   | || :|    | |||| ||  |  :::| |: :|| | |
Db         12 LVQAIFNGDPDEVRALIFKKEDVNFQD--NEKRTPLHAAAYLGDAEIIELLILSGARVNA 69

Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGKIDVCIVLLQ 219
              :|   | ||| | :    | | :||:| || |||| ||  |||| ||
Db         70 KDSKWLTPLHRAVASCSEEAVQVLLKHSADVNARDKNWQ-TPLHIAA------------ 115

Qy        220 HGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCH 279
                                 : |||                |   | ||  | :
Db        116 ---------------------ANKAV--------------------KCAEALVPLLSNVN 134

Qy        280 ASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVK 339
              ||    | || ||     ::|:|||  ||:::| || |   :| |   || || :|||
Db        135 VSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFDKKDRRAIHWAAYMGHIEVVKLLVS 194

Qy        340 HGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKER 399
              ||| |   |   :|||| |||   : |   ||  | |     | :  : : :|
Db        195 HGAEVTCKDKKSYTPLHAAASSGMISVVKYLLDLGVDMNEPNAYGNTPLHVAC------- 247

Qy        400 LAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETALHCAAASPYPKRKQIC- 457
                : |  ::          : : :    :|| |: :   | || |||| :     :|
Db        248 ----YNGQDVV----------VNELIDCGAIVNQKNEKGF-TPLHFAAASTH---GALCL 289

Qy        458 ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 517
              |||:  ||::| |:|:  ||||: :         : :::  | ::  |  | | || ||
Db        290 ELLVGNGADVNMKSKDGKTPLHMTALHGRFSRSQTIIQSGAVIDCEDKNGNTPLHIAARY 349

Qy        518 GHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQQLLQEGISLGN-SEADRQLL 571
              ||      |:: | |     : |    | :     ::  ::|| |  :     :  | |
Db        350 GHELLINTLITSGADTAKRGIHGMFPLHLAALSGFSDCCRKLLSSGFDIDTPDDFGRTCL 409

Qy        572 EAAKA-GDVETVKKL-----------------------CTVQ----------SVNCRDI 596
              || | |::| :  |                       |  |          |||  |:
Db        410 HAAAAGGNLECLNLLLNTGADFNKKDKFGRSPLHYAAANCNYQCLFALVGSGASVN--DL 467

Qy        597 EGRQSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 655
              : |  ||||:||  :      :|||:: |:   :|| |   :| : :|||    :|:
Db        468 DERGCTPLHYAATSDTDGKCLEYLLRNDANPGIRDKQGYNAVHYSAAYGHRLCLQLIASE 527

Qy        656 GAV-----VNVADLWK-------FTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
                :     :  |:              :||| ||  | ::  ::|:|    |   :|  | |||
Db        528 TPLDVLMETSGTDMLSDSDNRATISPLHLAAYHGHHQALEVLVQSLLDLDVRNSSGRTPL 587

Qy        704 DLVK-DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAG 762
              ||   |  :  |:|   | :   |  | |                          ||:| ||
Db        588 DLAAFKGHVECVDVLINQGASI-LVKDYILKR-------------------TPIHAAAT 626

Qy        763 YNNLEVAEYLLQHGADVNA---QDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 819
                : |    |: :    ||   ||  |  ||  :   || |    |:   | |:| |||
Db        627 NGHSECLRLLIGNAEPQNAVDIQDGNGQTPLMLSVLNGHTDCVYSLLNKGANVDAKDKWG 686

Qy        820 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA----DDVSALLTAAMPPS 875
              | ||  |  |  :    || |||   |::  |:||: | :|      : ||| :|
Db        687 RTALHRGAVTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAACGHIGVLGALLQSAASMD 746

Qy        876 ALP-------------SCYKPQVLNG-------------VRSPGATADALSSGPSSPSS 908
              | |                :||     ||                ::  |    |    : :
Db        747 ANPATADNHGYTALHWACY-----NGHETCVELLLEQEVFQKTEGNAFSPLHCAVINDNE 801

Qy        909 LSAASSLDNLSGSFSELSSVVSSSGTEGASSL----------------EKKEVPGVDFS 951
              :|   :| |        :|:|::: ::| : |                :|  ||
Db        802 GAAEMLIDTLG------ASIVNATDSKGRTPLHAAAFTDHVECLQLLLSHNAQVNSVD-- 853

Qy        952 ITQFVRNLGLEHLMDIFEREQI-TLDVLVEMGHKEL 986
                  : |   ||   |  |  |:::||     ||
Db        854 ------STGKTPLMMAAENGQTNTVEMLVSSASAEL 883


RESULT 33
US-10-291-172-343
; Sequence 343, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc
;  TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-045
;  CURRENT APPLICATION NUMBER: US/10/291,172
;  CURRENT FILING DATE:  2000-11-08
;  PRIOR APPLICATION NUMBER: 09/693,267
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 09/665,363
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: 09/616,847
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 09/596,193
;  PRIOR FILING DATE: 2000-06-17
;  PRIOR APPLICATION NUMBER: 09/574,454
;  PRIOR FILING DATE: 2000-05-19
;  PRIOR APPLICATION NUMBER: 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 343
;   LENGTH: 1053
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-291-172-343

  Query Match           11.8%;  Score 760.5;  DB 15;  Length 1053;
  Best Local Similarity  27.6%;  Pred. No. 2.4e-53;
  Matches  275;  Conservative  129;  Mismatches  357;  Indels  235;  Gaps   33;

Qy        102 LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
              | :|  ||| : |: |:   | || :|    | |||| ||  |  :::| |: :|| | |
Db         12 LVQAIFNGDPDEVRALIFKKEDVNFQD--NEKRTPLHAAAYLGDAEIIELLILSGARVNA 69

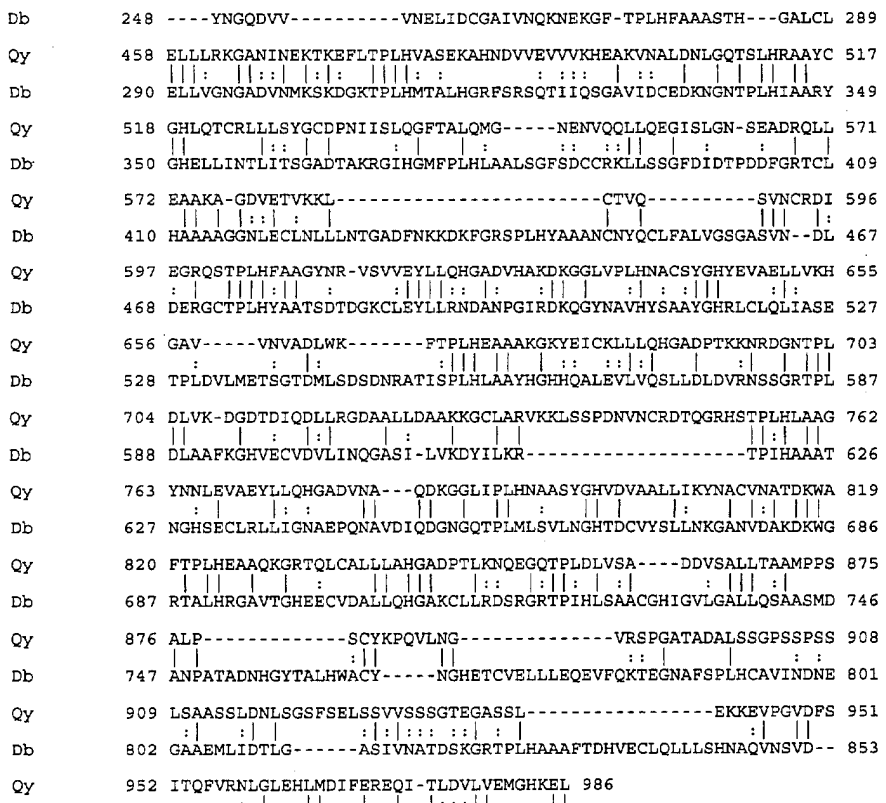
Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGKIDVCIVLLQ 219
              :|   | ||| | :    | | :||:| || |||| ||  |||| ||
Db         70 KDSKWLTPLHRAVASCSEEAVQVLLKHSADVNARDKNWQ-TPLHIAA------------ 115

Qy        220 HGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCH 279
                                 : |||                |   | ||  | :
Db        116 ---------------------ANKAV--------------------KCAEALVPLLSNVN 134

Qy        280 ASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVK 339
              ||    | || ||     ::|:|||  ||:::| || |   :| |   || || :|||
Db        135 VSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFDKKDRRAIHWAAYMGHIEVVKLLVS 194

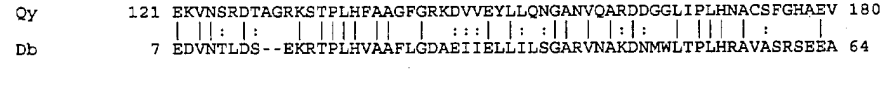
Qy        340 HGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKER 399
              ||| |   |   :|||| |||   : |   ||  | |     | :  : : :|
Db        195 HGAEVTCKDKKSYTPLHAAASSGMISVVKYLLDLGVDMNEPNAYGNTPLHVAC------- 247

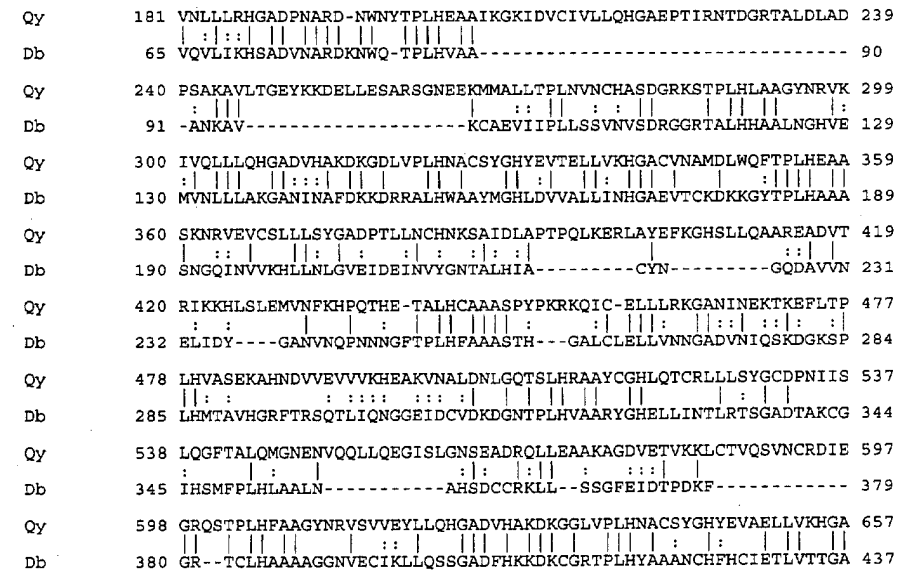
Qy        400 LAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETALHCAAASPYPKRKQIC- 457
                : |  ::          : : :    :|| |: :   | || |||| :     :|
```



RESULT 34
US-10-108-260A-4122
; Sequence 4122, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4122
; LENGTH: 919
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4122

Query Match 11.1%; Score 714.5; DB 15; Length 919;
Best Local Similarity 28.7%; Pred. No. 1.2e-49;
Matches 246; Conservative 109; Mismatches 287; Indels 215; Gaps 27;





RESULT 35
US-10-199-937-42
; Sequence 42, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582

```
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
;    LENGTH: 119
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;    OTHER INFORMATION: specified as protein-coding is vector sequence
US-10-199-937-42

  Query Match             10.1%;  Score 650;  DB 14;  Length 119;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-46;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1102 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 60

Qy       1162 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 119


RESULT 36
US-10-199-937-44
; Sequence 44, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
;    LENGTH: 119
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;    OTHER INFORMATION: specified as protein-coding is vector sequence
US-10-199-937-44

  Query Match             10.1%;  Score 650;  DB 14;  Length 119;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-46;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1102 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 60

Qy       1162 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 119


RESULT 37
US-10-199-937-8
; Sequence 8, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;    LENGTH: 118
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: SITE
;    LOCATION: (67)
;    OTHER INFORMATION: Xaa= unknown
US-10-199-937-8

  Query Match             9.9%;  Score 639;  DB 14;  Length 118;
  Best Local Similarity   99.2%;  Pred. No. 7.2e-45;
  Matches  117;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1103 LFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS 1162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS 60

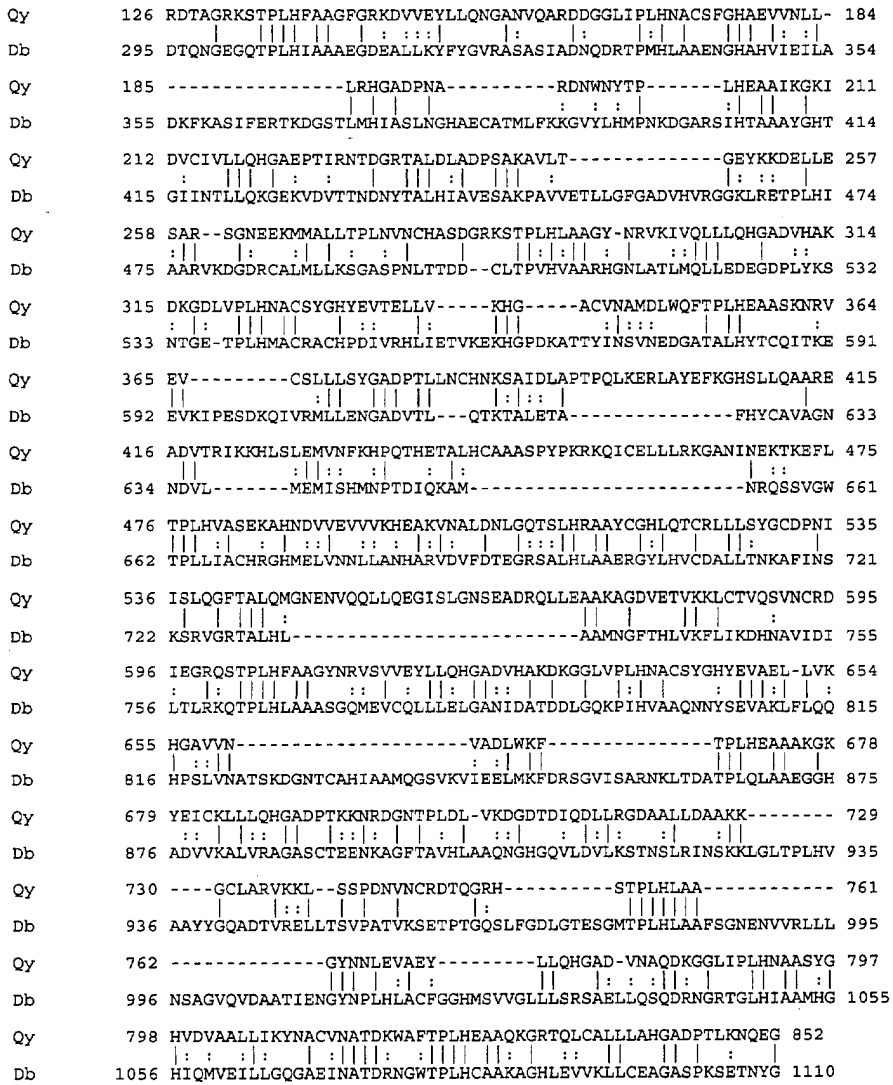
Qy       1163 CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CYICHRXLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 118


RESULT 38
US-10-369-978-4
; Sequence 4, Application US/10369978
; Publication No. US20030152991A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Walker, Richard G.
;  APPLICANT:  Willingham, Aarron
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
;  FILE REFERENCE: 02307E-097600US
;  CURRENT APPLICATION NUMBER: US/10/369,978
;  CURRENT FILING DATE:  2003-02-18
;  PRIOR APPLICATION NUMBER: US/09/392,812A
;  PRIOR FILING DATE: 1999-09-09
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;    LENGTH: 1619
;    TYPE: PRT
;    ORGANISM: Drosophila melanogaster
;    FEATURE:
;    OTHER INFORMATION: amino acid sequence derived from nompC cDNA
;    OTHER INFORMATION: sequence
US-10-369-978-4

  Query Match             9.8%;  Score 633.5;  DB 14;  Length 1619;
  Best Local Similarity   25.6%;  Pred. No. 1.7e-42;
  Matches  260;  Conservative  133;  Mismatches  341;  Indels  281;  Gaps   33;

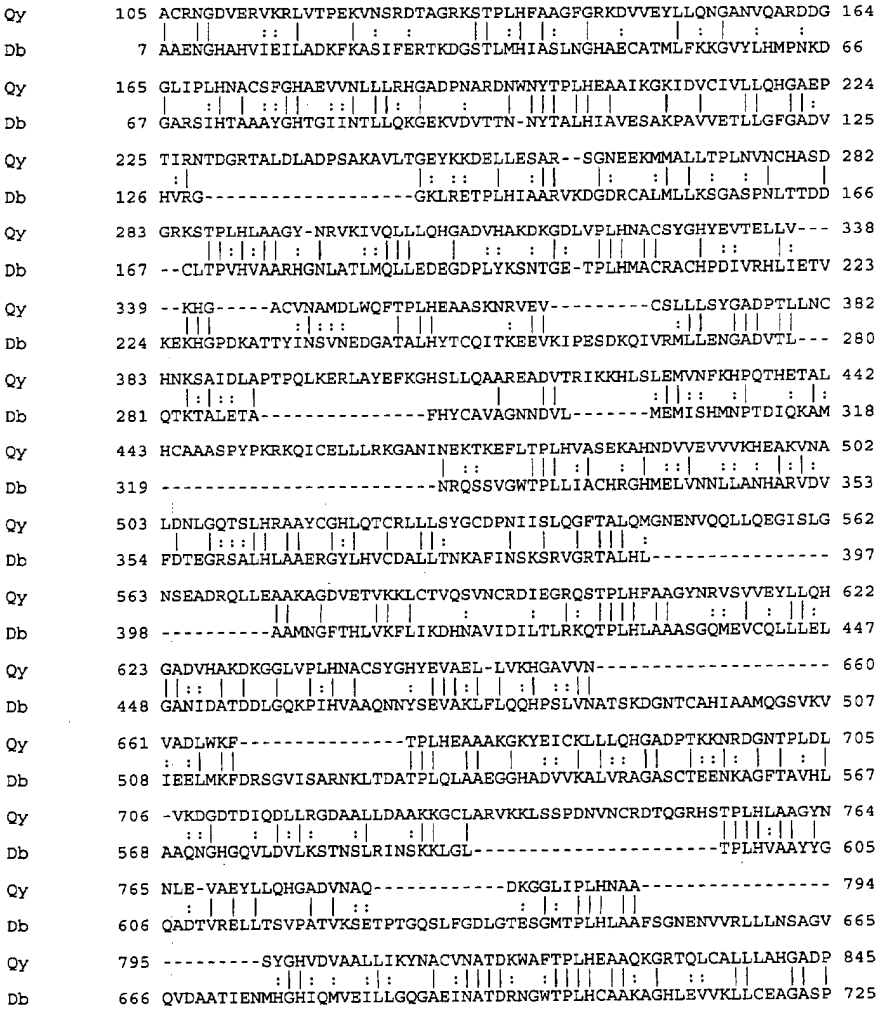
Qy         29 DPVTAG---SQAARALSASSPGG-----LALLLAGPGLLLRLLA-------LLLAVAAAR 73
              || : |   || |  | :|   |     |  |||  |  :|| |       |||||
Db        186 DPFSTGGSRSQTAVHLVSSRQTGTATNILRALLAAAGKDIRLKADGRGKIPLLLAVE--- 242

Qy         74 IMSGRRCAGGGAAC----ASAAAEAVEPAARE----LFEACRNGDVERVKRLVTPEKVNS 125
                    :|  : |    |:  || ::      |  | |  ||  | |::     |
Db        243 -------SGNQSMCRELLAAQTAEQLKATTANGDTALHLAARRRDVHMV-RILVDYGTNV 294
```



RESULT 39
US-10-369-978-2
; Sequence 2, Application US/10369978
; Publication No. US20030152991A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/10/369,978
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US/09/392,812A
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-10-369-978-2

Query Match 8.7%; Score 563; DB 14; Length 1704;
Best Local Similarity 24.8%; Pred. No. 1.3e-36;
Matches 210; Conservative 121; Mismatches 296; Indels 220; Gaps 22;



```
Qy        846 TLKNQEG 852
              :   |
Db        726 KSETNYG 732
```

RESULT 40
US-10-369-978-6
; Sequence 6, Application US/10369978
; Publication No. US20030152991A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/10/369,978
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US/09/392,812A
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1709
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-10-369-978-6

Query Match 8.4%; Score 545.5; DB 14; Length 1709;
Best Local Similarity 24.5%; Pred. No. 3.8e-35;
Matches 260; Conservative 114; Mismatches 352; Indels 335; Gaps 38;

```
Qy         64 ALLLAVAAAR-------IMSGRRC-----AGGGA---ACASAAAEAVEPAARELFEACRN 108
              | |||| | :       |  | |       |  |   |  :  | :|   :   |  | |
Db        154 AFLLAVKAGKDQIVDKMIRKGARVDYSTKDGRNATHIAAMYSGVETLELILKRYSELLRK 213

Qy        109 GDVERVKRLVTPEKVNSRDTAG-RKSTPLHFAAGFGRK---DVVEYLLQNGANVQARDDG 164
              |                     || :|   :| |      |    :|: :|::     | |
Db        214 G------------------AGPKKQLAIHVACERKSKKAFPIVKRILEDTDQRMAEDGD 254

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNAR--DNWNYTPLHEAAIKGKIDVCIVLLQHGA 222
              | :|:| |  ||:  :| |||   :|   |  |     | || ||  | |:     :  |
Db        255 GSLPIHLAFKFGNVNIVELLLSGPSDEQTRKADGNGDTLLHLAARSGNIEAVRTAIAAGC 314

Qy        223 E-PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHAS 281
              :   ::|  ||| |           :|| |             :: |: ::  |  : :
Db        315 DNANVQNRVGRTPLHEC-----LTVTGTQK-----GYVAEVGDQNMLKIMFKLRADANIH 364

Qy        282 DGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAK------------------------- 314
              |    ||:|:||       :|: |: : |  : |:
Db        365 DKEDKTPVHVAAERGDTSMVESLIDKFGGSIRARTRDGSTLLHIAACSGHTSTALAFLKR 424

Qy        315 ------DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCS 368
                   :|   : ||:| : |  :| ::|:  |  |:      :| || |    :  |
Db        425 VPLFMPNKKGALGLHSAAAAGFNDVVKMLIARGTNVDVRTRDNYTALHVAVQSGKASVVE 484

Qy        369 LLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE 428
              ||  |||                              ||  |:
Db        485 TLLGSGAD-------------------------IHVKGGELMDG---------------- 503

Qy        429 MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHND 488
                  || || || |      | |  ||| : |:    :|   ||| ||::  : :
Db        504 ----------ETCLHIAARS---GNKDIM-LLLDENADSKISSKIGETPLQVAAKSCNFE 549

Qy        489 VVEVVVKH----------EAKVNALDNLGQTSLHRAAYCGHLQ---------TCRLLLSY 529
                 :::||          :  ||   | | |:|| ||      |              ||: |
Db        550 AASMILKHLSEVLTQEQLKEHVNHRTNDGFTALHYAAEIEQRQLHFPGEDAKLVNLLIDY 609

Qy        530 GCDPNIISLQ-GFTALQM----GNENVQQLLQEGISLG------NSEADR---QLLEAAK 575
```

```
Db        610 GGMVEMPSLNANETAMHMAARSGNQAVLLAMVNKIGAGAVQIVQNKQSKNGWSPLLEACA 669

Qy        576 AGD------VETVKKLC---------------------------------------------- 586
              |       :  | ||
Db        670 RGHSGVANILLKVLVLCVGPGPGPGPRLQGRGYWTRTRARVTVPWLQYQGYWARTRTRTR 729

Qy        587 ----------------------TV--------QSVNCRDIEGRQSTPLHFAAGYNRVSV 615
                               ||          ::  |  ||  | || ||    :|:
Db        730 ARATGPGLQDQGYWARTRTRTKVTVPRLLGDHHARIDVFDEMGR--TALHLAAFNGHLSL 787

Qy        616 VEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVK-HGAVVNVADLWKFTPLHEAA 674
              |  |||| | |::| | |   ||| |  :|| :|  :||: ||| :     |  || ||
Db        788 VHLLLQHKAFVNSKSKTGEAPLHLAAQHGHVKVVNVLVQDHGAALEAITLDNQTALHFAA 847

Qy        675 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD-TDIQDLL-------RGDAALLD- 725
                |:  : : ||  ||:|  ::  | ||| |  : |  |:  |           |    :|
Db        848 KFGQLAVSQTLLALGANPNARDDKGQTPLHLAAENDFPDVVKLFLKMRNNNRSVLTAIDH 907

Qy        726 --------AAKKGCLARVKKLSSPDN--VNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQH 775
                     || || || |::|    |    |    |:   :| ||:||    :  : : ||::
Db        908 NGFTCAHIAAMKGSLAVVRELMMIDKPMVIQAKTKTLEATTLHMAAAGGHANIVKILLEN 967

Qy        776 GADVNAQD--------------------------------KGGLIPLHNAASYGHVDVA 802
              ||  ||:|                                | ||  || || ||: |
Db        968 GA--NAEDENSGMTALHLGAKNGFISILEAFDKILWKRCSRKTGLNALHIAAFYGNSDFV 1025

Qy        803 ALLIK------------YNACVNA--TDKWAFTPLHEAAQKGRTQLCALLLAHG--ADPT 846
              ::|              ||  ||   : :: ||||| ||  |    |  :|| |    | |
Db       1026 NEMLKHVQATVRSEPPIYNHHVNKEFSTEYGFTPLHLAAHSGHDSLVRMLLNQGVQVDAT 1085

Qy        847 ----LKNQEGQTPLDLVSAD------DVSALLTAAMPPSAL 877
                  :  :| :   ||::       |  |:||    :||
Db       1086 STTMMSEKEKERAKDLLNVAVFSENMAVELLITATEYNAAL 1126
```

RESULT 41
US-10-094-749-1924
; Sequence 1924, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1924
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-094-749-1924

Query Match 8.3%; Score 534; DB 15; Length 747;
Best Local Similarity 23.4%; Pred. No. 8.5e-35;
Matches 222; Conservative 112; Mismatches 292; Indels 322; Gaps 23;

```
Qy        104 EACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
              || : | :| || |:     |  |     | | :|    |  || : ||| ||:::   :
Db          2 EAAQEGHLELVKYLIA-AGANVHATTATGDTALTYACENGHTDVADVLLQAGADLEHESE 60

Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPN-ARDNWNYTPLHEAAIKGKIDVCIVLLQHGA 222
              ||  ||  |   ||   |  |:  ||: |    | ::| |  |   | : |  :|| |||
Db         61 GGRTPLMKAARAGHVCTVQFLISKGANVNRTTANNDHTVLSLACAGGHLAVVELLLAHGA 120

Qy        223 EPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALL---------- 271
              :|| |  || |                        |:|:|: |:   :  ||
Db        121 DPTHRLKDGSTM-------------------LIEAAKGGHTSVVCYLLDYPNNLLSAPP 160

Qy        272 ----------------------------------------TPLNVNCHASDGRKS---- 286
                                                     | | :   |  || |
Db        161 PDVTQLTPPSHDLNRAPRVPVQALPMVVPPQEPDKPPANVATTLPIRNKAVSGRASAMSN 220

Qy        287 TPLH-LAAGYNRVKI----------VQLLLQHGADVHAKDKGDL-VPLHNACSYGHYEVT 334
              || | :||  :: :           | :    |: |: : :    |  ||: || |:
Db        221 TPTHSIAASISQPQTPTPSPIISPSAMLPIYPAIDIDAQTESNHDTALTLACAGGHEELV 280

Qy        335 ELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTP 394
              : |:: || :   |  ||||  ||:   | |  :||  |||
Db        281 QTLLERGASIEHRDKKGFTPLILAATAGHVGVVEILLDNGAD----------------- 322

Qy        395 QLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRK 454
                        ::|  |                      :| :| |  | :     |:
Db        323 --------------IEAQSE-------------------RTKDTPLSLACSG---GRQ 344

Qy        455 QICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNAL--DNLGQTSLH 512
              :: |||| :|||   :     ||| :|:   : :::::::   |::|:     || : |
Db        345 EVVELLLARGANKEHRNVSDYTPLSLAASGGYVNIIKILLNAGAEINSRTGSKLGISPLM 404

Qy        513 RAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 572
              ||  ||    :|||  | | |
Db        405 LAAMNGHTAAVKLLLDMGSDIN---------------------------------- 426

Qy        573 AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 632
                                 ||  ::| | |     |  ||  ||   |:|  : |
Db        427 --------------------AQIETNRNTALTLACFQGRTEVVSLLLDRKANVEHRAKT 465

Qy        633 GLVPLHNACSYGHYEVAELLVKHGAVVNVADL--WKFTPLHEAAAKGKYEICKLLLQHGA 690
              || ||  | | |: ||  :|:  || ||    :  : | |  || || |: |:||:  ||
Db        466 GLTPLMEAASGGYAEVGRVLLDKGADVNAPPVPSSRDTALTIAADKGHYKFCELLIGRGA 525

Qy        691 DPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQ 750
                :|: ||
Db        526 HIDVRNKKGN---------------------------------------------- 535

Qy        751 GRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK--- 807
                  ||| |||   :|:| : |:| ||||:|  |   : ||  |   ||| |   :|
Db        536 ----TPLWLAANGGHLDVVQLLVQAGADVDAADNRKITPLMAAFRKGHVKVVRYLVKEVN 591

Qy        808 ----------YNACVNATDKWAFTPLH-----------EAA---------------QKG 830
                       | |:  |||          |           :||                :|
Db        592 QFPSDSECMRYIATI--TDKEMLKKCHLCMESIVQAKDRQAAEANKNASILLEELDLEKL 649

Qy        831 RTQLCALLLAHGADP-----TLKNQEGQTPLDLVSADDVSAL-LTAAMPPSALPSCYKPQ 884
              | :   | ||   :         | :| :  |: : | :       | ||      |   :|:
Db        650 REESRRLALAAKREKRKEKRRKKKEEQRRKLEEIEAKNKENFELQAAQEKEKLKVEDEPE 709

Qy        885 VLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSS 932
              ||     | ||         : :::  :::   |:||  : :: ::::
Db        710 VL--TEPPSAT---------TTTTIGISATWTTLAGSHGKRNNTITTT 746
```

RESULT 42
US-10-334-143-45

; Sequence 45, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-45

Query Match 8.2%; Score 528.5; DB 15; Length 3913;
Best Local Similarity 27.8%; Pred. No. 4e-33;
Matches 157; Conservative 74; Mismatches 184; Indels 149; Gaps 11;

```
Qy        439 ETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEA 498
              ||||| || |    : ::   |:: || :  | |:  ||||:::     |:|: :::  |
Db          3 ETALHMAARS---GQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGA 59

Qy        499 KVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEG 558
              ||     | || :|  ||       || :|    :| : :||
Db         60 SPNAATTSGYTPLHLSAREGHEDVAAFLLDHGASLSITTKKGF----------------- 102

Qy        559 ISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEY 618
                                                  |||| || | :: |
Db        103 ------------------------------------------TPLHVAAKYGKLEVANL 119

Qy        619 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 678
              |||  |   |  | || ||| |  | : :|| ||:  ||  : |    :|||| || | :
Db        120 LLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQ 179

Qy        679 YEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKK 737
                :|   ||::|||      | |   : |  ::|  |:  || | |
Db        180 MDIATTLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNA--------------- 224

Qy        738 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 797
                  |||  :  |  |||||||  : : ||| |:  || |:|| | |  |||     ||
Db        225 -----NVNLSNKSG--LTPLHLAAQEDRVNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYG 277

Qy        798 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 857
                :: :   |::::| ||| |  :||||:|||:| | :  :|| : | |       | | |
Db        278 NIKIVNFLLQHSAKVNAKTKNGYTPLHQAAQQGHTHIINVLLQNNASPNELTVNGNTALG 337

Qy        858 L------VSADDVSALLT-AAMPPSALPSCYK---PQVLNGVRSPGATADALSSGPSSPS 907
                    :|  |   ::|   |  : :   :|   |: :| |                |
Db        338 IARRLGYISVVDTLKIVTEETMTTTTVTEKHKMNVPETMNEVLD-----------MSDD 385

Qy        908 SLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDI 967
              :  |:: : ||                                        |:: |:
Db        386 EVRKANAPEMLSDG---------------------------------------EYISDV 405

Qy        968 FEREQ-ITLDVLVEMGHKELKEIG 990
              | |  :| |    :| ::|||:|
Db        406 EEGEDAMTGDTDKYLGPQDLKELG 429
```

RESULT 43
US-10-182-243-56
; Sequence 56, Application US/10182243
; Publication No. US20040048310A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD

; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: MARTINEZ, RICARDO
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
; TITLE OF INVENTION: ENZYMES
; FILE REFERENCE: 038602/1366
; CURRENT APPLICATION NUMBER: US/10/182,243
; CURRENT FILING DATE: 2003-07-07
; PRIOR APPLICATION NUMBER: PCT/US01/02337
; PRIOR FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 765
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-243-56

Query Match 7.4%; Score 478; DB 12; Length 765;
Best Local Similarity 24.5%; Pred. No. 3.9e-30;
Matches 160; Conservative 79; Mismatches 225; Indels 190; Gaps 10;

```
Qy        58 LLLRLLALLLAVAAARIMSGRRCAGGGAACASA---AAEAVEPAARELFEACRNGDVERV 114
             :|| ||   :|| :: :: :      :|  :     |  | ::|| ::     ::|
Db       287 ILLSLLQSRVAVPESKALARK------VSCKLSLRQPGEVNEDISQELMDSDSGNYLKRA 340

Qy       115 KRLVTPEKVNSRDTA----GRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLH 170
              :|   : :  ||        | |||||    | : |  || :  :|  :   |  ||
Db       341 LQLSDRKNLVPRDEELCIYENKVTPLHFLVAQGSVEQVRLLLAHEVDVDCQTASGYTPLL 400

Qy       171 NACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTD 230
              |      ::  ||| |||| |  |   : ||| ||  |      :|| |||    :  :
Db       401 IAAQDQQPDLCALLLAHGADANRVDEDGWAPLHFAAQNGDDGTARLLLDHGACVDAQERE 460

Qy       231 GRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT--PLNVNCHASDGRKSTP 288
             | | | |                 | | : ||  : | | ::|:  ||
Db       461 GWTPLHL--------------------AAQNNFENVARLLVSRQADPNLHEAEGK--TP 497

Qy       289 LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD 348
             ||:|| :  | :|:||   ||:: |: :    ||| |   |      : |:| ||  :|:|
Db       498 LHVAAYFGHVSLVKLLTSQGAELDAQQRNLRTPLHLAVERGKVRAIQHLLKSGAVPDALD 557

Qy       349 LWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS 408
              : ||| ||:: :  :| :|| |||
Db       558 QSGYGPLHTAAARGKYLICKMLLRYGA------------------------------- 584

Qy       409 LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN 468
                          |||:
Db       585 ----------------SLEL---------------------------------------- 588

Qy       469 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 528
              | :  ||||:|: | | ::: :: :   | : ||  :  | || ||  |          ||
Db       589 -PTHQGWTPLHLAAYKGHLEIIHLLAESHANMGALGAVNWTPLHLAARHGEEAVVSALLQ 647

Qy       529 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 588
              | |||     |:
Db       648 CGADPNAAEQSGW------------------------------------------------ 660

Qy       589 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 648
                        |||| |   :    |  ||:| |:|||::| |  | | |  |:  :
Db       661 ------------TPLHLAVQRSTFLSVINLLEHHANVHARNKVGWTPAHLAALKGNTAI 707

Qy       649 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTP 702
              ::||: || ::| |     |||  |    |  |   | :|:     |: |
Db       708 LKVLVEAGAQLDVQDGVSCTPLQLALRSRKQGIMSFL--EGKEPSVATLGGSKP 759
```

RESULT 44
US-10-104-047-2529
; Sequence 2529, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2529
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2529

Query Match 7.3%; Score 472.5; DB 15; Length 657;
Best Local Similarity 26.5%; Pred. No. 8.7e-30;
Matches 173; Conservative 69; Mismatches 217; Indels 195; Gaps 13;

```
Qy       123 VNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVN 182
             ||| |          |          |:| :  || |||:|    |   |  | ||   | :|||
Db       189 VNSED----DRTSCIVRQALEREDSIRTLLDNGASVNQCDSNGRTLLANAAYSGSLDVVN 244

Qy       183 LLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSA 242
             ||:  |||    |  :|||  || :|   |   |:  ||      : || |||  |
Db       245 LLVSRGADLEIEDAHGHTPLTLAARQGHTKVVNCLIGCGANINHTDQDGWTALRSA---- 300

Qy       243 KAVLTGEYKKDELLESARSGNEEKMMALL-TPLNVNCHASDGRKSTPLHLAAGYNRVKIV 301
                                |  |: | : |||   : |:|  :| |  | |  ||      ||
Db       301 ----------------AWGGHTEVVSALLYAGVKVDCADADSR--TALRAAAWGGHEDIV 342

Qy       302 QLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASK 361
              ||||||:|:  |       |  |   || |: | |: ||| ||  |:   | |  ||
Db       343 LNLLQHGAEVNKADNEGRTALIAAAYMGHREIVEHLLDHGAEVNHEDVDGRTALSVAA-- 400

Qy       362 NRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRI 421
                :|                          |          |||:
Db       401 ----LC-----------------------VPAS----------KGHA------------ 410

Qy       422 KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVA 481
                                                :  ||: :|| ::    |: :||| ||
Db       411 --------------------------------SVVSLLIDRGAEVDHCDKDGMTPLLVA 437

Qy       482 SEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGF 541
             : : | |||:::::  | |:  :| |:| |  ||  ||       || :|
Db       438 AYEGHVDVVDLLLEGGADVDHTENNGRTPLLAAASMGHASVVNTLLFWGA---------- 487

Qy       542 TALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQS 601
                                                          :|:  | |||
Db       488 ------------------------------------------------AVDSIDSEGR-- 497

Qy       602 TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNV 661
             | |  |:    | ||  ||  | | : ::   |  |||  |    ||  : | ||::  ||  |
Db       498 TVLSIASAQGNVEVVRTLLDRGLDENHREDAGWTPLHMAAFEGHRLICEALIEQGARTNE 557

Qy       662 ADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK-DGDTDIQDLLRGD 720
              |      |  |: :| |:  ::||:: ::  ::  ||   | :   :|  || :||
Db       558 IDNDGRIPFILASQEGHYDCVQILLENKSNIDQRGYDGRNALRVAALEGHRDIVELLFSH 617

Qy       721 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ 774
              |                :|||:|  || :  |:: |  | | :|||  :
Db       618 GA------------------DVNCKDADGRPT--LYILALENQLTMAEYFFR 649
```

RESULT 45
US-10-128-174-3
; Sequence 3, Application US/10128174
; Publication No. US20030199462A1
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
; FILE REFERENCE: UM-06967
; CURRENT APPLICATION NUMBER: US/10/128,174
; CURRENT FILING DATE: 2002-04-23

```
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 765
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-128-174-3

  Query Match             7.2%;  Score 465;  DB 14;  Length 765;
  Best Local Similarity   30.1%;  Pred. No. 4.7e-29;
  Matches  144;  Conservative   65;  Mismatches  181;  Indels   88;  Gaps    8;

Qy        385 KSAIDLAP----TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440
              | |: |:      |: :|   || |   |     :  | :::  |:           ||
Db        338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVEQVRLLLA----------HEV 386

Qy        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500
               : |  || |                          ||| :|::    |:  ::: | |
Db        387 DVDCQTASGY-----------------------TPLLIAAQDQQPDLCALLLAHGADA 421

Qy        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560
              | :|   |   || ||  |  :| |||| :|   :    :|:| | :  :|
Db        422 NRVDEDGWAPLHFAAQNGDDRTARLLLDHGACVDAQEREGWTPLHLAAQN---------- 471

Qy        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620
              | |   :|| : :|              | |: ||:  |||| || :  ||:|: |
Db        472 --NFENVARLLVSRQA------------DPNLREAEGK--TPLHVAAYFGHVSLVKLLT 514

Qy        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680
               ||:: |: :    ||| |   |     : |:| ||| :  |   : ||| |||:|||
Db        515 SQGAELDAQQRNLRTPLHLAVERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYL 574

Qy        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740
              |||:||::||        | ||| |                  || || |  :  | :
Db        575 ICKMLLRYGASLELPTHQGWTPLHL-------------------AAYKGHL-EIIHLLA 613

Qy        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800
               : |       : |||||| :    |   ||| ||| || :: |  ||| |
Db        614 ESHANMGALGAVNWTPLHLAARHGEEAVVSALLQCGADPNAAEQSGWTPLHLAVQRSTFL 673

Qy        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 858
                |::::| |:| :|  :|| | || || | :  :|:  ||   :::    ||| |
Db        674 SVINLLEHHANVHARNKVGWTPAHLAALKGNTAILKVLVEAGAQLDVQDGVSCTPLQL 731


Search completed: March 25, 2004, 16:54:49
Job time : 66.812 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:38:10 ; Search time 23.3162 Seconds
(without alignments)
5115.640 Million cell updates/sec

Title: US-09-843-159B-4
Perfect score: 6464
Sequence: 1 RCSARRGAAGGQGAQRGARV.....AYPEYLITYQIMRPEGMVDG 1240

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	895	13.8	1765	2	T42714	ankyrin 3, splice
2	895	13.8	1961	2	T42716	ankyrin 3, splice
3	892.5	13.8	4377	2	A55575	ankyrin 3, long sp
4	891	13.8	1943	2	T42713	ankyrin 3, splice
5	889	13.8	1940	2	T42715	ankyrin 3, splice
6	880.5	13.6	1856	2	B35049	ankyrin 1, erythro
7	880.5	13.6	1880	2	A35049	ankyrin 1, erythro
8	878.5	13.6	1881	1	SJHUK	ankyrin 1, erythro
9	877	13.6	3924	2	S37431	ankyrin 2, neurona
10	876	13.6	1848	2	S37771	ankyrin, erythrocy
11	874	13.5	1862	2	I49502	ankyrin - mouse
12	866.5	13.4	1549	2	T13940	ankyrin - fruit fl
13	858	13.3	2039	2	T15347	ankyrin-related un
14	657	10.2	791	2	T42691	hypothetical prote
15	563.5	8.7	934	1	H71274	probable ankyrin -
16	552	8.5	1411	2	S30355	alpha-latroinsecto
17	507.5	7.9	1058	2	D82654	ankyrin-like prote
18	494.5	7.7	1062	2	T30255	inversin - mouse
19	489.5	7.6	1062	2	T14151	Inv protein - mous
20	485	7.5	1401	2	S11527	alpha-latrotoxin p
21	463	7.2	1188	2	T19552	hypothetical prote
22	452.5	7.0	426	2	AE2149	hypothetical prote
23	426	6.6	991	2	T25412	hypothetical prote
24	416.5	6.4	1184	2	T00253	gene Ankhzn protei
25	399	6.2	2584	2	T24158	hypothetical prote
26	399	6.2	2606	2	T24157	hypothetical prote
27	389.5	6.0	397	2	T46445	hypothetical prote
28	370	5.7	1031	2	T43458	hypothetical prote
29	350	5.4	557	2	T46507	hypothetical prote
30	348.5	5.4	606	2	AC2508	hypothetical prote
31	345	5.3	1423	1	I37275	death-associated p
32	343	5.3	1107	2	T21280	hypothetical prote
33	329.5	5.1	1435	2	T32930	hypothetical prote
34	321	5.0	1398	2	T21884	hypothetical prote
35	315.5	4.9	656	2	A34793	sex-determining pr
36	307	4.7	1004	2	A55142	myosin-light-chain
37	306	4.7	1016	2	T19006	ankyrin related pr
38	306	4.7	1322	2	A59288	myosin heavy chain
39	305	4.7	679	2	B45771	2-5A-dependent RNA
40	299.5	4.6	633	2	T27499	hypothetical prote
41	299	4.6	247	2	D84448	probable ankyrin [
42	297	4.6	1001	2	S30385	G9a protein - huma
43	297	4.6	1295	2	T21720	hypothetical prote
44	294.5	4.6	900	2	A42024	transcription fact
45	285.5	4.4	658	2	S68418	protein phosphatas

ALIGNMENTS

RESULT 1
T42714
ankyrin 3, splice form 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42714
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene f
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 1587/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 13.8%; Score 895; DB 2; Length 1765;
Best Local Similarity 30.3%; Pred. No. 1.7e-46;
Matches 294; Conservative 134; Mismatches 409; Indels 134; Gaps 25;

```
Qy         68 AVAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRD 127
              |  | |   |::     :   :| |  :| |     :  :||                ||  :
Db         10 AKPAHRKRKGKKSDANASYLRAARAGHLEKA----LDYIKNG-----------VDVNICN 54

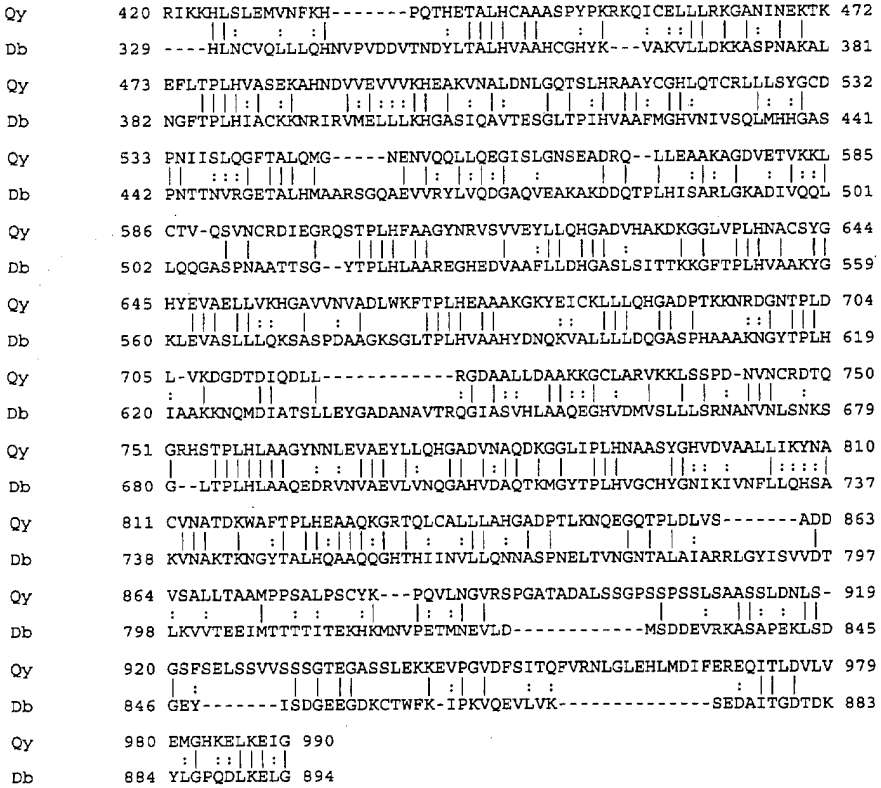
Qy        128 TAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRH 187
                |  :  || |:  |  :||  |||  |||  |     |   || |   | |||| :|: :
Db         55 QNGLNA--LHLASKEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTN 112

Qy        188 GADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSA 242
              ||: ||:    :|||: || :  ::|   || :||  ::   || | | :|     |
Db        113 GANVNAQSQNGFTPLYMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVV 172

Qy        243 KAVLTGEYK---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVK 299
               :|  : |   :   |  |   :: |  |||   : |          ||||:|| |  :
Db        173 SLLLENDTKGKVRLPALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNIN 232

Qy        300 IVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAA 359
              :  |||   | |    : |: ||| |   |:  : :||:  || ::|       ||||  |
Db        233 VATLLLNRAAAVDFTARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGA 292

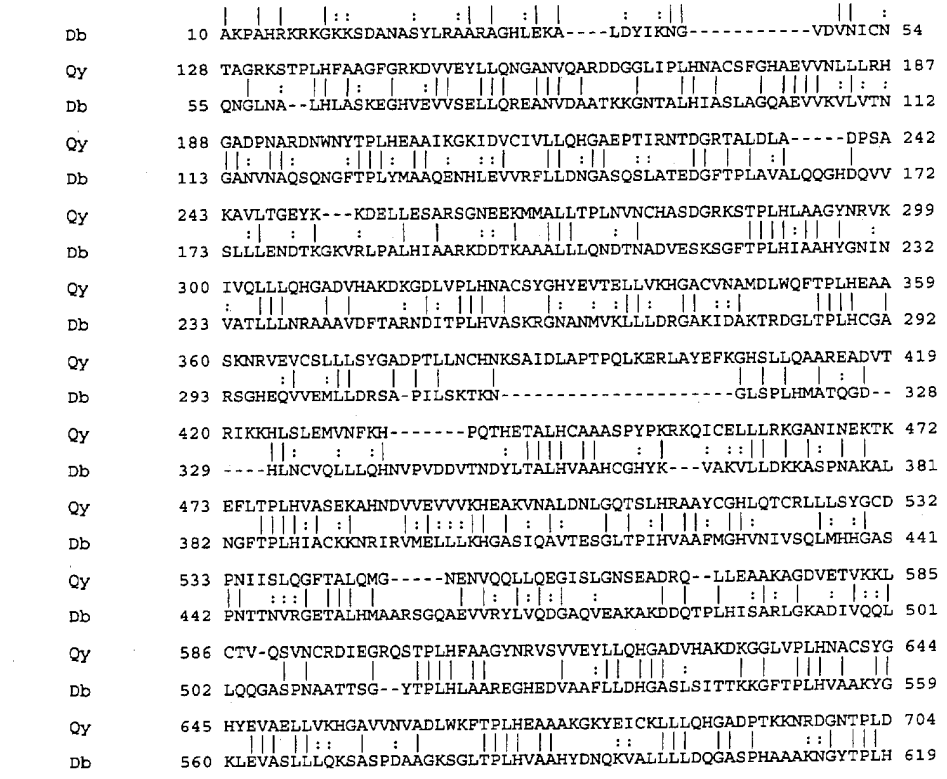
Qy        360 SKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVT 419
                :|  :||   | | |   |                    | | |  | : |
Db        293 RSGHEQVVEMLLDRSA-PILSKTKN-------------------GLSPLHMATQGD-- 328
```



RESULT 2
T42716
ankyrin 3, splice form 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42716
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42716
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1961 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 13.8%; Score 895; DB 2; Length 1961;
Best Local Similarity 30.3%; Pred. No. 2e-46;
Matches 294; Conservative 134; Mismatches 409; Indels 134; Gaps 25;





RESULT 3
A55575
ankyrin 3, long splice form - human
N;Alternate names: ankyrin G
C;Species: Homo sapiens (man)

```
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C;Accession: A55575
R;Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A;Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax
A;Reference number: A55575; MUID:95138209; PMID:7836469
A;Accession: A55575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4377 <KOR>
A;Cross-references: GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025
C;Genetics:
A;Gene: GDB:ANK3
A;Cross-references: GDB:424503; OMIM:600465
A;Map position: 10q21-10q21
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: alternative splicing; peripheral membrane protein
F;73-105/Domain: ankyrin repeat homology <AN01>
F;106-138/Domain: ankyrin repeat homology <AN02>
F;139-171/Domain: ankyrin repeat homology <AN03>
F;172-200/Domain: ankyrin repeat homology <AN04>
F;201-233/Domain: ankyrin repeat homology <AN05>
F;234-266/Domain: ankyrin repeat homology <AN06>
F;267-299/Domain: ankyrin repeat homology <AN07>
F;300-332/Domain: ankyrin repeat homology <AN08>
F;333-365/Domain: ankyrin repeat homology <AN09>
F;366-398/Domain: ankyrin repeat homology <AN10>
F;399-431/Domain: ankyrin repeat homology <AN11>
F;432-464/Domain: ankyrin repeat homology <AN12>
F;465-497/Domain: ankyrin repeat homology <AN13>
F;498-530/Domain: ankyrin repeat homology <AN14>
F;531-563/Domain: ankyrin repeat homology <AN15>
F;564-596/Domain: ankyrin repeat homology <AN16>
F;597-629/Domain: ankyrin repeat homology <AN17>
F;630-662/Domain: ankyrin repeat homology <AN18>
F;663-695/Domain: ankyrin repeat homology <AN19>
F;696-728/Domain: ankyrin repeat homology <AN20>
F;729-761/Domain: ankyrin repeat homology <AN21>
F;762-794/Domain: ankyrin repeat homology <AN22>
F;795-827/Domain: ankyrin repeat homology <AN23>

  Query Match             13.8%;  Score 892.5;  DB 2;  Length 4377;
  Best Local Similarity   30.7%;  Pred. No. 1e-45;
  Matches  277;  Conservative  126;  Mismatches  365;  Indels  135;  Gaps   21;

Qy        136 LHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD 195
              || |:  |  :||  |||  ||| |    |   || |   | |||| :|: :||: ||:
Db         78 LHLASKEGHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQS 137

Qy        196 NWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEY 250
                :|||: || :  ::|   || :||  ::   || | | :|     |     :| :
Db        138 QNGFTPLYMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDT 197

Qy        251 K---KDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQH 307
              |   :   |  |   :: |  |||   : |         ||||:|| |  : :  |||
Db        198 KGKVRLPALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINVATLLLNR 257

Qy        308 GADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVC 367
               | |    : |: ||| |   |:  : :||:  || ::|      ||||  |     :|
Db        258 AAAVDFTARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVV 317

Qy        368 SLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL 427
              :||   | | |    |                  | | |  | : |       ||:
Db        318 EMLLDRAA-PILSKTKN-------------------GLSPLHMATQGD------HLNC 349

Qy        428 EMVNFKH-------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 480
              :  :|          : |||| ||   : |   : ::|| | || | |     ||||:
Db        350 VQLLLQHNVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTPLHI 406

Qy        481 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 540
              | :|    |:|:::|| | : |:   | | :| ||: ||:     |: :|  ||  :::|
Db        407 ACKKNRIKVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRG 466

Qy        541 FTALQMG-----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVN 592
              |||  |        | |: |:|:|  :      | |  |  :|: |  : |::|      | |
Db        467 ETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGASPN 526

Qy        593 CRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELL 652
                  |  |||| :|      |  :|| ||| :    | |  ||| |  || ||| ||
Db        527 AATTSG--YTPLHLSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVANLL 584

Qy        653 VKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDT 711
              ::  |  : |      |||| ||     ::  ||| || |    ::| ||| :  |
Db        585 LQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQM 644

Qy        712 DIQDLL-----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLH 758
              ||   |           :| ||:  ||::| :  |  |   : |||  :  |   ||||
Db        645 DIATTLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNANVNLSNKSG--LTPLH 702

Qy        759 LAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
              |||  : : ||| |:  || |:|| | |  |||    ||:: :   |::::| |||  |
Db        703 LAAQEDRVNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKN 762

Qy        819 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLT-AA 871
               :||||:|||:| | :  :|| : | |      | | | :      :|  |   ::|
Db        763 GYTPLHQAAQQGHTHIINVLLQNNASPNELTVNGNTALGIARRLGYISVVDTLKIVTEET 822

Qy        872 MPPSALPSCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSV 928
              |  : :   :|   |: :| |                |   :  |:: : ||
Db        823 MTTTTVTEKHKMNVPETMNEVLD-----------MSDDEVRKANAPEMLSDG------- 863

Qy        929 VSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ-ITLDVLVEMGHKELK 987
                                                |:: |: | |  :| |    :| ::||
Db        864 --------------------------------EYISDVEEGEDAMTGDTDKYLGPQDLK 890

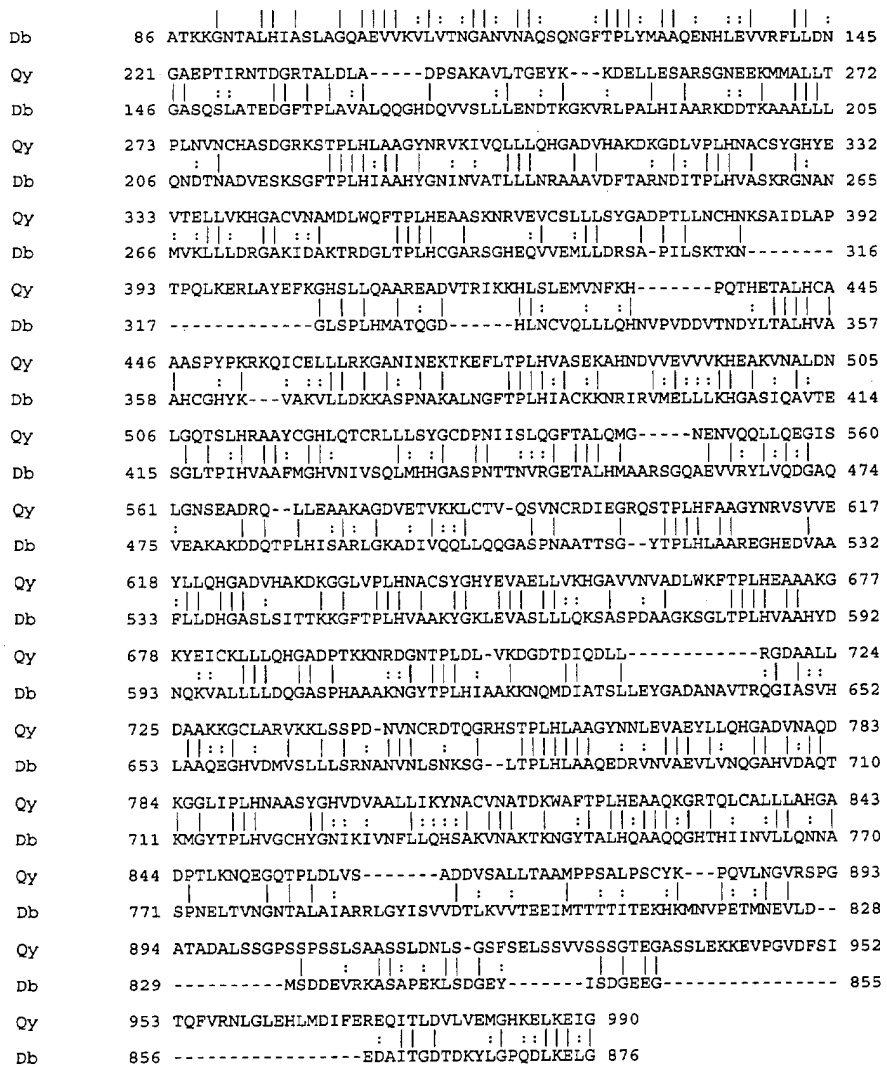
Qy        988 EIG 990
              |:|
Db        891 ELG 893


RESULT 4
T42713
ankyrin 3, splice form 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42713
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene f
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42713
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1943 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 855/1
C;Function:
A;Description: supposed to play an important role in the polarized distribution of many i
A;Note: major kidney ankyrin
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

  Query Match             13.8%;  Score 891;  DB 2;  Length 1943;
  Best Local Similarity   30.6%;  Pred. No. 3.5e-46;
  Matches  287;  Conservative  127;  Mismatches  380;  Indels  144;  Gaps   24;

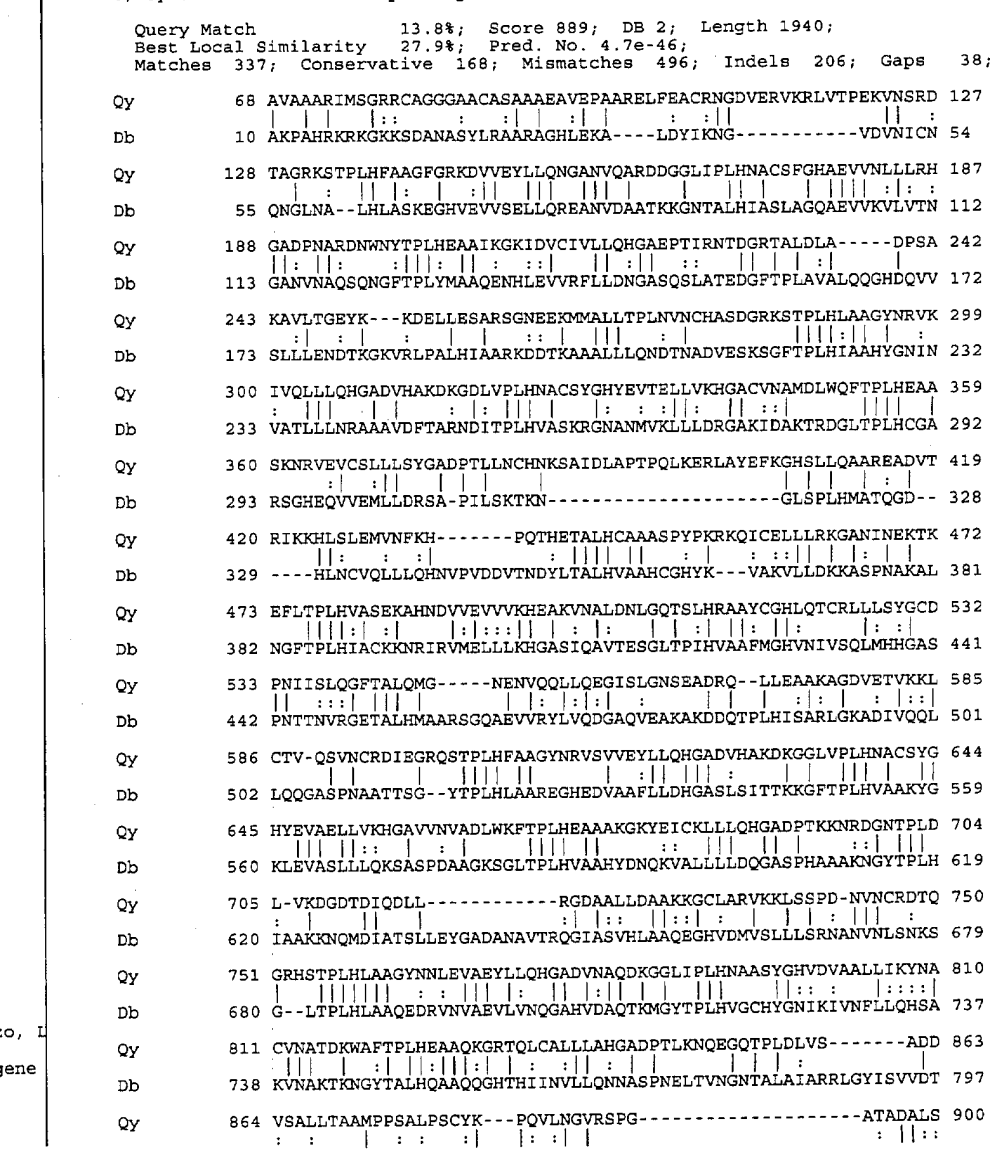
Qy        105 ACRNGDVER----VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQA 160
              | | | :|:    :|  |    ||  :  |  :  || |:  |  :||  |||  ||| |
Db         31 AARAGHLEKALDYIKNGV---DVNICNQNGLNA--LHLASKEGHVEVVSELLQREANVDA 85

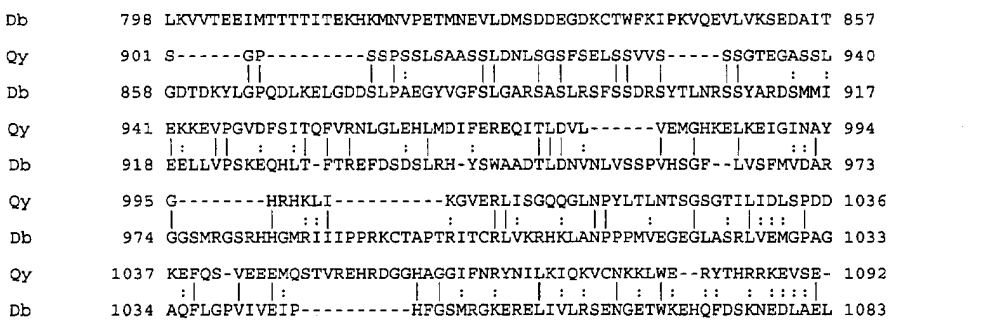
Qy        161 RDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQH 220
```



RESULT 5
T42715
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 834/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 13.8%; Score 889; DB 2; Length 1940;
Best Local Similarity 27.9%; Pred. No. 4.7e-46;
Matches 337; Conservative 168; Mismatches 496; Indels 206; Gaps 38;

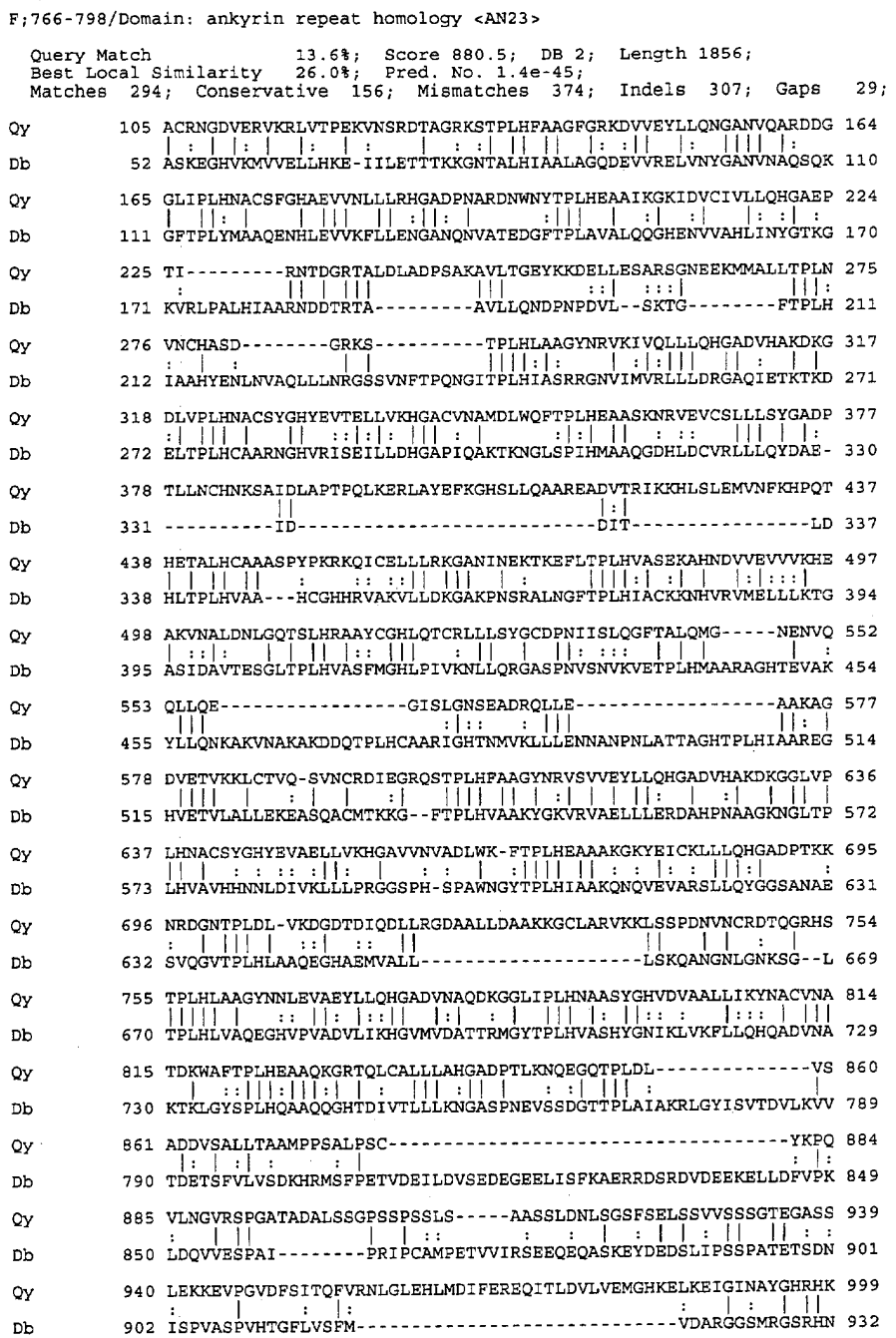




RESULT 6
B35049
ankyrin 1, erythrocyte splice form 3 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C;Accession: B35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:90175370; PMID:1689849
A;Accession: B35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1856 <LAM>
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>

F;766-798/Domain: ankyrin repeat homology <AN23>

Query Match 13.6%; Score 880.5; DB 2; Length 1856;
Best Local Similarity 26.0%; Pred. No. 1.4e-45;
Matches 294; Conservative 156; Mismatches 374; Indels 307; Gaps 29;



```
Qy       1000 LIKGVE-------------RLISGQQGLNPYLTLNTSGSGTILIDLSP------------ 1034
              :: |                ||:  |:   |        |  : :| | |
Db        933 GLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEGLASRIIALGPTGAQFLSPVIVE 992

Qy       1035 ---------DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 1076
                      |:|   :   |   |  :|||          :||    : :: |
Db        993 IPHFASHGRGDRELVVLRSENGSVWKEHR---------SRYGESYLDQILN 1034


RESULT 7
A35049
ankyrin 1, erythrocyte splice form 2 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C;Accession: A35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:90175370; PMID:1689849
A;Accession: A35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1880 <LAM>
A;Cross-references: GB:M28880
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; cytoskeleton
F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

  Query Match             13.6%;  Score 880.5;  DB 2;  Length 1880;
  Best Local Similarity   26.0%;  Pred. No. 1.5e-45;
  Matches  294;  Conservative  156;  Mismatches  374;  Indels  307;  Gaps   29;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
              | : | |: |  |:  | :    | : :| || ||  |: :||  |:  |||| |:
Db         52 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 110

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
              |  ||: |    | |||  || :||: |      :|||  |  :|  :|   |: :| :
Db        111 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 170

Qy        225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
              :         || | |||         |||       ::|  :::|         |||:
```

```
Db        171 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 211

Qy        276 VNCHASD--------GRKS----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
              :  |  :          | |          ||||:|:    | :|:|||  || :  | |
Db        212 IAAHYENLNVAQLLLNRGSSVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETKTKD 271

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
              :| ||| |   ||  ::|:|: ||| : |     :|:| ||  : ::   ||| | |:
Db        272 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYDAE- 330

Qy        378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437
                       ||                        |:|
Db        331 ----------ID-------------------------DIT----------------LD 337

Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
              | | || ||   :   :: ::|| |||  | :     ||||:| :|  |  |:|:::|
Db        338 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRVMELLLKTG 394

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQ 552
              | ::|:   | | || |:: |||   : ||  |  ||: :::  | | |       | :
Db        395 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 454

Qy        553 QLLQE----------------GISLGNSEADRQLLE----------------AAKAG 577
              |||                    :|::   : |||                ||: |
Db        455 YLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAREG 514

Qy        578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
              |||| |  : | |  :|  |||| || | :| | | ||:  |  :|  | || |
Db        515 HVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVRVAELLLERDAHPNAAGKNGLTP 572

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
              || |  : : :: :||:  |   : :  | :|||| || : : |: : |||:|  :
Db        573 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQVEVARSLLQYGGSANAE 631

Qy        696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
              : | ||| |  ::|  ::  ||                      ||  | |  :  |
Db        632 SVQGVTPLHLAAQEGHAEMVALL-------------------LSKQANGNLGNKSG--L 669

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
              ||||| |   :: ||: |::||  |:|  : |  ||| |: ||:: :   |::: | |||
Db        670 TPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 729

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL-------------VS 860
              |  ::|||:|||:| | :  ||| :|| |   : :| |||  :                |
Db        730 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDGTTPLAIAKRLGYISVTDVLKVV 789

Qy        861 ADDVSALLTAAMPPSALPSC-----------------------------------YKPQ 884
              |: | :| :    : |                                       : |:
Db        790 TDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISFKAERRDSRDVDEEKELLDFVPK 849

Qy        885 VLNGVRSPGATADALSSGPSSPSSLS-----AASSLDNLSGSFSELSSVVSSSGTEGASS 939
              :   | ||          | | ::      :   :  | : | | : || || : :
Db        850 LDQVVESPAI--------PRIPCAMPETVVIRSEEQEQASKEYDEDSLIPSSPATETSDN 901

Qy        940 LEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHK 999
              :      |    :  |:                                :   | :  | ||
Db        902 ISPVASPVHTGFLVSFM----------------------------VDARGGSMRGSRHN 932

Qy       1000 LIKGVE-------------RLISGQQGLNPYLTLNTSGSGTILIDLSP------------ 1034
              :: |                ||:  |:   |        |  : :| | |
Db        933 GLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEGLASRIIALGPTGAQFLSPVIVE 992

Qy       1035 ---------DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 1076
                      |:|   :   |   |  :|||          :||    : :: |
Db        993 IPHFASHGRGDRELVVLRSENGSVWKEHR---------SRYGESYLDQILN 1034


RESULT 8
SJHUK
ankyrin 1, erythrocyte splice form 1 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2
C;Species: Homo sapiens (man)
```

```
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S08275; A33219; PC2220; A35443
R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A;Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure w
A;Reference number: S08275; MUID:90158830; PMID:2137557
A;Accession: S08275
A;Molecule type: mRNA
A;Residues: 1-1881 <LU1>
A;Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A;Accession: A33219
A;Molecule type: protein
A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,'X'
X',1367;1383-1427;1601-1630;1686-1698,'D',1700;1763-1772 <LUX>
A;Note: 845-Arg and 1392-Thr were also found
R;Hermann, J.; Barel, M.; Frade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membrane
A;Reference number: PC2220; MUID:95071348; PMID:7526850
A;Accession: PC2220
A;Molecule type: protein
A;Residues: 910-929 <HER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchanger
A;Reference number: A35443; MUID:90285190; PMID:2141335
A;Accession: A35443
A;Molecule type: protein
A;Residues: 'X',5,'X',7-12;403-417,'X',419-422,'H',424,'LQ';797-800,'L',802-814;862-863,
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; phosphoprotein
F;2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F;2-1512,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>
F;2-827/Domain: 89K #status predicted <DOM1>
F;2-827/Region: anion exchange protein binding
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>
F;828-1382/Domain: 62K #status predicted <DOM2>
F;828-1382/Region: spectrin binding
F;1383-1881/Domain: 55K #status predicted <DOM3>

  Query Match             13.6%;  Score 878.5;  DB 1;  Length 1881;
  Best Local Similarity   25.8%;  Pred. No. 2e-45;
  Matches  292;  Conservative  157;  Mismatches  375;  Indels  307;  Gaps   28;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
Db         52 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 110

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
Db        111 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 170

Qy        225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
Db        171 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 211

Qy        276 VNCHASDGRKS-----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
Db        212 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETKTKD 271

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
Db        272 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYDAE- 330

Qy        378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437
Db        331 ----------ID-------------------------DIT---------------LD 337

Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
Db        338 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHVRVMELLLKTG 394

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQ 552
Db        395 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 454

Qy        553 QLLQE----------------GISLGNSEADRQLLE-----------------AAKAG 577
Db        455 YLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLENNANPNLATTAGHTPLHIAAREG 514

Qy        578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
Db        515 HVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVRVAELLLERDAHPNAAGKNGLTP 572

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
Db        573 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQVEVARSLLQYGGSANAE 631

Qy        696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
Db        632 SVQGVTPLHLAAQEGHAEMVALL-------------------LSKQANGNLGNKSG--L 669

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
Db        670 TPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 729

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL-------------VS 860
Db        730 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDGTTPLAIAKRLGYISVTDVLKVV 789

Qy        861 ADDVSALLTAAMPPSALPSC-----------------------------------YKPQ 884
Db        790 TDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISFKAERRDSRDVDEEKELLDFVPK 849

Qy        885 VLNGVRSPGATADALSSGPSSPSSLS-----AASSLDNLSGSFSELSSVVSSSGTEGASS 939
Db        850 LDQVVESPAI--------PRIPCAMPETVVIRSEEQEQASKEYDEDSLIPSSPATETSDN 901

Qy        940 LEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHK 999
Db        902 ISPVASPVHTGFLVSFM----------------------------VDARGGSMRGSRHN 932

Qy       1000 LIKGVE-------------RLISGQQGLNPYLTLNTSGSGTILIDLSP------------ 1034
Db        933 GLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEGLASRIIALGPTGAQFLSPVIVE 992

Qy       1035 ---------DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 1076
Db        993 IPHFASHGRGDRELVVLRSENGSVWKEHR---------SRYGESYLDQILN 1034
```

RESULT 9
S37431
ankyrin 2, neuronal long splice form - human
N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid
N;Contains: ankyrin 2, short form
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C;Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
R;Chan, W.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A;Reference number: A39643; MUID:91302466; PMID:1830053
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2077 <OT1>
A;Cross-references: GB:X56957
A;Accession: B39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443,3585-3924 <OTT>
A;Cross-references: EMBL:X56958
R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
Genomics 10, 858-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92009921; PMID:1833308
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 463-474,'PE',477-495 <TSE>
A;Cross-references: GB:M37123; NID:g178647; PIDN:AAA62828.1; PID:g178648
R;Chan, W.; Kordeli, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A;Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and se
A;Reference number: A49462; MUID:94075409; PMID:8253844
A;Accession: A49462
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3924 <RES>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
C;Genetics:
A;Gene: GDB:ANK2
A;Cross-references: GDB:127607; OMIM:106410
A;Map position: 4q25-4q27
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F;2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F;63-95/Domain: ankyrin repeat homology <AN01>
F;96-128/Domain: ankyrin repeat homology <AN02>
F;129-161/Domain: ankyrin repeat homology <AN03>
F;162-190/Domain: ankyrin repeat homology <AN04>
F;191-223/Domain: ankyrin repeat homology <AN05>
F;232-264/Domain: ankyrin repeat homology <AN06>
F;265-297/Domain: ankyrin repeat homology <AN07>
F;298-330/Domain: ankyrin repeat homology <AN08>
F;331-363/Domain: ankyrin repeat homology <AN09>
F;364-396/Domain: ankyrin repeat homology <AN10>
F;397-429/Domain: ankyrin repeat homology <AN11>
F;430-462/Domain: ankyrin repeat homology <AN12>
F;463-495/Domain: ankyrin repeat homology <AN13>
F;496-528/Domain: ankyrin repeat homology <AN14>
F;529-561/Domain: ankyrin repeat homology <AN15>
F;562-594/Domain: ankyrin repeat homology <AN16>
F;595-627/Domain: ankyrin repeat homology <AN17>
F;628-660/Domain: ankyrin repeat homology <AN18>
F;661-693/Domain: ankyrin repeat homology <AN19>
F;694-726/Domain: ankyrin repeat homology <AN20>
F;727-759/Domain: ankyrin repeat homology <AN21>
F;760-792/Domain: ankyrin repeat homology <AN22>
F;793-825/Domain: ankyrin repeat homology <AN23>

Query Match 13.6%; Score 877; DB 2; Length 3924;
Best Local Similarity 29.7%; Pred. No. 7.7e-45;
Matches 293; Conservative 137; Mismatches 402; Indels 156; Gaps 23;

```
Qy      105 ACRNGDVER-VKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDD 163
            | | |:::: |: |     :|: :  |  :  || ||   :|: ||  |::| :
Db       38 AARAGNLDKVVEYLKGGIDINTCNQNGLNA--LHLAAKEGHVGLVQELLGRGSSVDSATK 95

Qy      164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
              |  || |   | |||| :|:: ||: ||:    :|||: || :  |||   ||::||
Db       96 KGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPLYMAAQENHIDVVKYLLENGAN 155

Qy      224 PTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL----LESARSGNEEKMMALLTPLN 275
             :    || |    ||       | |:|      | ::     |  |   :: |  |||
Db      156 QSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLPALHIAARKDDTKSAALLLQ-- 213

Qy      276 VNCHASDGRKS----------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324
             | | :| :           ||||:|| |  | :  |||  || |    :  : |||
Db      214 -NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATLLLNRGAAVDFTARNGITPLHV 272

Qy      325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384
             |   |:  : :||:  |  ::|      |||| ||     :|  |||  || | |     |
Db      273 ASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGHDQVVELLLERGA-PLLARTKN 331

Qy      385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHP-----QTHE 439
                                  | | |  | :  |  |||    :  | |       :
Db      332 --------------------GLSPLHMAAQGDHVECVKHL----LQHKAPVDDVTLDYL 366

Qy      440 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 499
             |||| ||    :    :: :||| | || | :     ||||:| :|    |:|::||: |
Db      367 TALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPLHIACKKNRIKVMELLVKYGAS 423

Qy      500 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN-----ENVQQL 554
             : |:   | | :| ||: |||     |||  |  |:: :::| ||| |        | |: |
Db      424 IQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNIRGETALHMAARAGQVEVVRCL 483

Qy      555 LQEGI----------------SLGNSEADRQLLE----------------AAKAGDV 579
             |: |                || :|  : ||:                  :|: | |
Db      484 LRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTNGYTPLHISAREGQV 543

Qy      580 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 639
             :   |   : :    : :  |||| || |  : | : |||  |   :  | || |||
Db      544 DVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHV 602

Qy      640 ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDG 699
             | | : :|| ||::  ||  :      :|||| || | : :|   || | :||:      : |
Db      603 AAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQMQIASTLLNYGAETNIVTKQG 662

Qy      700 NTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLH 758
             |||  |  ::| ||:  ||    | :  : |  |                  | ||
Db      663 VTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----------------------TSLH 700

Qy      759 LAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW 818
             |||  : : ||: | :|||| :|  | |  ||  |  ||:| :   |:|  | |||  |
Db      701 LAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYGNVKMVNFLLKQGANVNAKTKN 760

Qy      819 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLTAAM 872
             :||||:|||:|  | :  :|| ||| |       | | :      :|  |   ::|  :
Db      761 GYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALAIAKRLGYISVVDTLKVVTEEV 820

Qy      873 PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSEL------- 925
              :     |   : :  |    : |           :|     | ::|   |
Db      821 TTTTTTITEK----HKLNVPETMTEVL--------DVSDEEGDDTMTGDGGEYLRPEDLK 868

Qy      926 ----SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RNLGLEHLMDIFEREQITLDVLV 979
                |: ||   :| : |      |   |:  |   |:  | |    : |:   :|  |
```

```
Db        869 ELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRSHTLSHAS--YLRDSAVMDDSV 926

Qy        980 EMGHKELKEIGINAYGHRHKLIKGVERL 1007
              :   ::  :   |  : ::|  | | |
Db        927 VIPSHQVSTLAKEAERNSYRLSWGTENL 954


RESULT 10
S37771
ankyrin, erythrocyte - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: S37771
R;Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found am
A;Reference number: S37771; MUID:93252825; PMID:8486643
A;Accession: S37771
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1848 <BIR>
A;Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;48-80/Domain: ankyrin repeat homology <AN01>
F;81-113/Domain: ankyrin repeat homology <AN02>
F;114-146/Domain: ankyrin repeat homology <AN03>
F;147-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;275-307/Domain: ankyrin repeat homology <AN08>
F;308-340/Domain: ankyrin repeat homology <AN09>
F;341-373/Domain: ankyrin repeat homology <AN10>
F;374-406/Domain: ankyrin repeat homology <AN11>
F;407-439/Domain: ankyrin repeat homology <AN12>
F;440-472/Domain: ankyrin repeat homology <AN13>
F;473-505/Domain: ankyrin repeat homology <AN14>
F;506-538/Domain: ankyrin repeat homology <AN15>
F;539-571/Domain: ankyrin repeat homology <AN16>
F;572-604/Domain: ankyrin repeat homology <AN17>
F;605-637/Domain: ankyrin repeat homology <AN18>
F;638-670/Domain: ankyrin repeat homology <AN19>
F;671-703/Domain: ankyrin repeat homology <AN20>
F;704-736/Domain: ankyrin repeat homology <AN21>
F;737-769/Domain: ankyrin repeat homology <AN22>
F;770-802/Domain: ankyrin repeat homology <AN23>

  Query Match             13.6%;  Score 876;  DB 2;  Length 1848;
  Best Local Similarity   29.7%;  Pred. No. 2.7e-45;
  Matches  250;  Conservative  132;  Mismatches  281;  Indels  178;  Gaps   21;

Qy        105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
              | : | |: |  |:  | :    |  : :| || ||  |: :||  |:  |||| |:
Db         56 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 114

Qy        165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
              |  ||: |    | |||  || :||: |      :|||  |  :|  :|   |: :| :
Db        115 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 174

Qy        225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
              :         || | |||          |||       ::|  :::|          |||:
Db        175 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 215

Qy        276 VNCHASDGRKS----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
              :  |  :  :                ||||:|:     | :|:|||  || :  : |
Db        216 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETRTKD 275

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
              :| ||| |   ||  ::|:|: ||| : |      :|:| ||  : ::   ||| | |:
Db        276 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYNAE- 334

Qy        378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437
```

```
                  ||                          |:|
Db        335 ----------ID--------------------------DIT----------------LD 341

Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
              | | || ||   :    :: ::|| ||| | :     ||||:| :| | |:|:::|
Db        342 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHIRVMELLLKTG 398

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM----GNENVQQ 553
              | ::|:   | | || |:: |||   : ||  |  ||: :::  | | |    |:  | :
Db        399 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 458

Qy        554 LLQEGISLGNSEAD-----------------RQLLE-----------------AAKAG 577
              | :  :  |::|                  : |||                  ||: |
Db        459 YLLQNKAKANAKADDQTPLHCAARIGHTGMVKLLLENGASPNLATTAGHTPLHTAAREG 518

Qy        578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
              |:|   |   : | |   :|   |||| || | :| : | ||:| |  :|  | || |
Db        519 HVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVRLAELLLEHDAHPNAAGKNGLTP 576

Qy        637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
              || |  : : :: :||:  |   : :  |  :|||| || : : |: : |||:|     :
Db        577 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQIEVARSLLQYGGSANAE 635

Qy        696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
              :  | ||| |  ::| |::   ||                ||   | |  :  |
Db        636 SVQGVTPLHLAAQEGHTEMVALL-------------------LSKQANGNLGNKSG--L 673

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
              ||||| :   :: ||: |::||  |:|  :  |  ||| |: ||:: :   |::: | |||
Db        674 TPLHLVSQEGHVPVADVLIKHGVTVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 733

Qy        815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALL 868
              |  ::|||:|||:| | :  ||| :|| |   :  | ||| :       :|  ||  ::
Db        734 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNGTTPLAIAKRLGYISVTDVLKVV 793

Qy        869 T 869
              |
Db        794 T 794


RESULT 11
I49502
ankyrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C;Accession: I49502
R;White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A;Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory doma
A;Reference number: I49502; MUID:92345717; PMID:1386265
A;Accession: I49502
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1862 <RES>
A;Cross-references: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C;Genetics:
A;Gene: Ank-1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;40-72/Domain: ankyrin repeat homology <AN01>
F;73-105/Domain: ankyrin repeat homology <AN02>
F;106-138/Domain: ankyrin repeat homology <AN03>
F;139-167/Domain: ankyrin repeat homology <AN04>
F;168-200/Domain: ankyrin repeat homology <AN05>
F;201-233/Domain: ankyrin repeat homology <AN06>
F;234-266/Domain: ankyrin repeat homology <AN07>
F;267-299/Domain: ankyrin repeat homology <AN08>
F;300-332/Domain: ankyrin repeat homology <AN09>
F;333-365/Domain: ankyrin repeat homology <AN10>
F;366-398/Domain: ankyrin repeat homology <AN11>
F;399-431/Domain: ankyrin repeat homology <AN12>
F;432-464/Domain: ankyrin repeat homology <AN13>
F;465-497/Domain: ankyrin repeat homology <AN14>
```

F;498-530/Domain: ankyrin repeat homology <AN15>
F;531-563/Domain: ankyrin repeat homology <AN16>
F;564-596/Domain: ankyrin repeat homology <AN17>
F;597-629/Domain: ankyrin repeat homology <AN18>
F;630-662/Domain: ankyrin repeat homology <AN19>
F;663-695/Domain: ankyrin repeat homology <AN20>
F;696-728/Domain: ankyrin repeat homology <AN21>
F;729-761/Domain: ankyrin repeat homology <AN22>
F;762-794/Domain: ankyrin repeat homology <AN23>

Query Match 13.5%; Score 874; DB 2; Length 1862;
Best Local Similarity 29.7%; Pred. No. 3.7e-45;
Matches 250; Conservative 132; Mismatches 281; Indels 178; Gaps 21;

```
Qy      105 ACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDG 164
Db       48 ASKEGHVKMVVELLHKE-IILETTTKKGNTALHIAALAGQDEVVRELVNYGANVNAQSQK 106

Qy      165 GLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 224
Db      107 GFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPLAVALQQGHENVVAHLINYGTKG 166

Qy      225 TI---------RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLN 275
Db      167 KVRLPALHIAARNDDTRTA---------AVLLQNDPNPDVL--SKTG--------FTPLH 207

Qy      276 VNCHASDGRKS-----------------TPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317
Db      208 IAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRRGNVIMVRLLLDRGAQIETRTKD 267

Qy      318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377
Db      268 ELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPIHMAAQGDHLDCVRLLLQYNAE- 326

Qy      378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437
Db      327 ----------ID-------------------------DIT---------------LD 333

Qy      438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497
Db      334 HLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGFTPLHIACKKNHIRVMELLLKTG 390

Qy      498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM----GNENVQQ 553
Db      391 ASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAK 450

Qy      554 LLQEGISLGNSEAD-----------------RQLLE-----------------AAKAG 577
Db      451 YLLQNKAKANAKAKDDQTPLHCAARIGHTGMVKLLLENGASPNLATTAGHTPLHTAAREG 510

Qy      578 DVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 636
Db      511 HVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVRLAELLLEHDAHPNAAGKNGLTP 568

Qy      637 LHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKGKYEICKLLLQHGADPTKK 695
Db      569 LHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQNQIEVARSLLQYGGSANAE 627

Qy      696 NRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
Db      628 SVQGVTPLHLAAQEGHTEMVALL-------------------LSKQANGNLGNKSG--L 665

Qy      755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA 814
Db      666 TPLHLVSQEGHVLVADVLIKHGVTVDATTRMGYTPLHVASHYGNIKLVKFLLQHQADVNA 725

Qy      815 TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALL 868
Db      726 KTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNGTTPLAIAKRLGYISVTDVLKVV 785

Qy      869 T 869
Db      786 T 786
```

RESULT 12
T13940
ankyrin - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C;Accession: T13940
R;Dubreuil, R.R.; Yu, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994
A;Title: Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Drosoph
A;Reference number: Z17820; MUID:95024098; PMID:7937942
A;Accession: T13940
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1549 <DUB>
A;Cross-references: EMBL:L35601; NID:g557083; PID:g557084; PIDN:AAC37208.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0011747

Query Match 13.4%; Score 866.5; DB 2; Length 1549;
Best Local Similarity 27.4%; Pred. No. 8e-45;
Matches 311; Conservative 173; Mismatches 492; Indels 159; Gaps 34;

```
Qy       98 AARELFEACRNGDVERVKRLV---TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQN 154
Db       38 ATISFLRAARSGDIKKVMDFLDCGEISDINSCNANGLNA--LHLAAKDGYVDICCELLRR 95

Qy      155 GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVC 214
Db       96 GIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCC 155

Qy      215 IVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDELLESARSGNEEK 266
Db      156 RTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLENDVRGKVRLPALHIAAKKNDVN 215

Qy      267 MMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNAC 326
Db      216 AAKLLLQHDPNADIVSKSGFTPLHIAAHYGNVDIATLLLNNKADVNYVAKHNITPLHVAC 275

Qy      327 SYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN-K 385
Db      276 KWGKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLLQQNA-PILTKTKNGL 334

Qy      386 SAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE----------MVNF 432
Db      335 SALHMAAQGE------HDEAAHLLLDNKAPVDEVTVDYLTALHVAAHCGHVKVAKLLLDY 388

Qy      433 K-HPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 487
Db      389 KANPNARALNGFTPLHIACKK---NRIKMVELLIKHGANIGATTESGLTPLHVASFMGCI 445

Qy      488 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL--- 544
Db      446 NIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDAIVREGQTPLHVA 504

Qy      545 -QMGNEN-VQQLLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600
Db      505 SRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENIVQVLLENGAEN-NAVTKKG 563

Qy      601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660
Db      564 FTPLHLACKYGKQNVVQILLQNGASIDFQGKNDVTPLHVATHYNNPSIVELLLKNGSSPN 623

Qy      661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRG 719
Db      624 LCARNGQCAIHIACKKNYLEIAMQLLQHGADVNIISKSGFSPLHLAAQGGNVDMVQLLLE 683

Qy      720 DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADV 779
Db      684 YGVISAAAKNGL----------------------TPLHVAAQEGHVLVSQILLEHGANI 720

Qy      780 NAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLL 839
```

```
Db        721 SERTRNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAAQQGHIMIINLLL 780

Qy        840 AHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG----------V 889

Db        781 RHKANPNALTKDGNTALHIASN------LGYVTVMESLKIVTSTSVINSNIGAIEEKLKV 834

Qy        890 RSPGATADALSSGPSSPSSLSAAS-------SLDNLSGSFSELSSVVSSSGTE----GAS 938

Db        835 MTPELMQETLLSDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFDTTNTDHDLTDVS 894

Qy        939 SLEKKEVPGVDFSITQFVRNLGLEHLMD--IFEREQITLDVLVEMGHKELKEIGINAYGH 996

Db        895 VLNKKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF---LVDARGGSMRGY 948

Qy        997 RHKLIKGVE------------RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS-- 1041

Db        949 RHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSRILEMSPVDGMFLSPI 1008

Qy       1042 -VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100

Db       1009 TLEVPHYGTLRK-----------NEREII-ILRSDNGESWREH-NLYKDIIGEDINQTEE 1055

Qy       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIG--GGT 1153

Db       1056 ---FHSDRIVR-IVTQNVPHFFAVV-------------SRVRQEVHVIGPDGGT 1092


RESULT 13
T15347
ankyrin-related unc-44 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Aug-2002
C;Accession: T15347; T15346; T15344; T15345; A57282; B57282; C57282
R;Gattung, S.
submitted to the EMBL Data Library, February 1996
A;Description: The sequence of C. elegans cosmid B0350.
A;Reference number: Z18332
A;Accession: T15347
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2039 <GAT>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208873; PIDN:AAA93443.1
A;Accession: T15346
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1000,'SKLQHRT',1002-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVA
32,'S',2034-2035,'GSPTRRSVEPEEHRHSQHEDHEGST' <GA2>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1
A;Accession: T15344
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1718,'KW',1903-1905,'NRLADESSPS',1916-1917,'QRSTIVAESTSEQVPE',1934-1935,'E
<GA3>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208874; PIDN:AAA93444.1
A;Accession: T15345
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVAES',1956-1957,'EQVPE
PTRRSVEPEEHRHSQHEDHEGST' <GA4>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208876; PIDN:AAA93446.1
R;Otsuka, A.J.; Franco, R.; Yang, B.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakulpoo
J. Cell Biol. 129, 1081-1092, 1995
A;Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in Cae
A;Reference number: A57282; MUID:95263663; PMID:7744957
A;Accession: A57282
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-852,'GGG',856-1000,'SKLQHRT',1002-1319,'IG',1322-1595,'DA',1598-1718,'KWEE
,'SHRED',2007-2008,'TI',2011,2017,'TT',2020-2022,'SHIS' <OTS>
A;Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608
A;Accession: B57282
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 831-852,'GGG',856-1319,'IG',1322-1595,'DA',1598-1718,'KWEELNRL',1727,1799,'E'
V',1945-1947,'VT',1950,'SH',1975,'SESP',1980-1981,'SPTRRSVEPEEHRHS',1984-1985,'EDHEGS',19
A;Cross-references: GB:U21731
A;Accession: C57282
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 194,'F',196,'I',198,'DC',201,'G',409,'AV',412,'Q',414,824,'S',826,'ER',829,'S
4,'TIV',1828,'ESTS',1833,'QV',1836,'E',1934-1935,'EQS',1939,'ESES',1944,'REDDGTIVTT',1945
A;Cross-references: GB:U21732; NID:g790603; PIDN:AAA85853.1; PID:g790604
C;Genetics:
A;Gene: CESP:unc-44
A;Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1979/
C;Superfamily: ankyrin; ankyrin repeat homology
F;164-192/Domain: ankyrin repeat homology <AN04>
F;358-390/Domain: ankyrin repeat homology <AN1>
F;391-423/Domain: ankyrin repeat homology <AN11>

  Query Match             13.3%;  Score 858;  DB 2;  Length 2039;
  Best Local Similarity   26.3%;  Pred. No. 4.1e-44;
  Matches  319;  Conservative  163;  Mismatches  443;  Indels  286;  Gaps   39;

Qy         91 AAEAVEPAARE----LFEACRNGDVERVKRLV-TPEKVNSRDTAGRKSTPLHFAAGFGRK 145

Db         22 APAAPEPGRAEGSASFLRAARAGDLEKVLELLRAGTDINTSNANGLNS--LHLASKEGHS 79

Qy        146 DVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 205

Db         80 EVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPLYMA 139

Qy        206 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 265

Db        140 AQENHEEVVKYLLKHGANQALSTEDGFTPLAVA-------------------LQQGHDR 179

Qy        266 KMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 324

Db        180 VVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTPLHI 233

Qy        325 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 384

Db        234 AAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IIDSRT 290

Qy        385 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHC 444

Db        291 KDLL-------------------------------------------------TPLHC 299

Qy        445 AAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN--A 502

Db        300 AARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVDDVT 356

Qy        503 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN------------ 550

Db        357 VDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELLLKY 414

Qy        551 ---VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 600

Db        415 RAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV---------------RG 458

Qy        601 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 660

Db        459 ETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGANSN 518

Qy        661 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLL-- 717

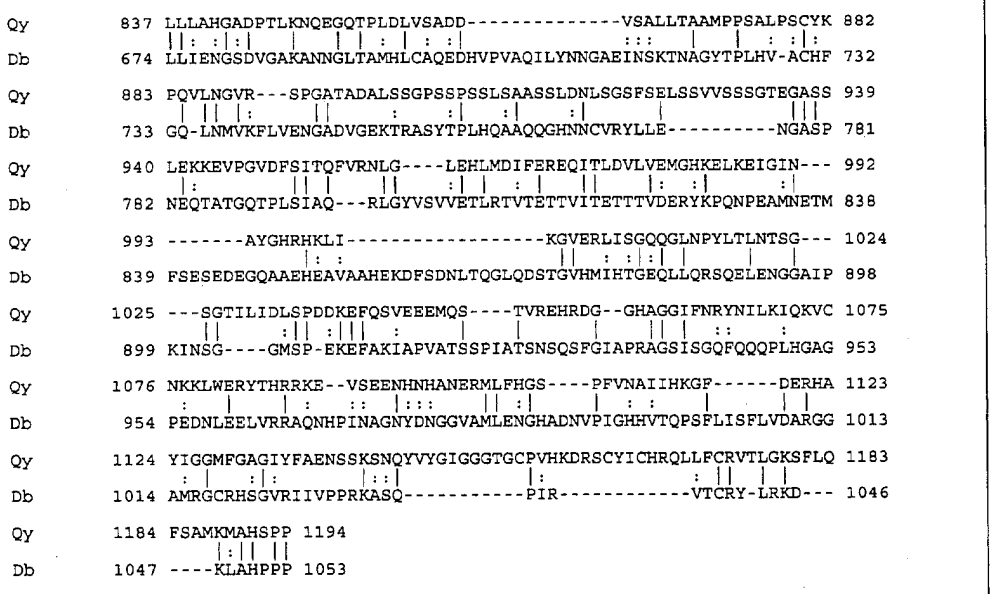
Db        519 ATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRLLLE 578

Qy        718 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHG 776

Db        579 RGTPV-----------------------DIEGKNQVTPLHVAAHYNNDKVAMLLLENG 613

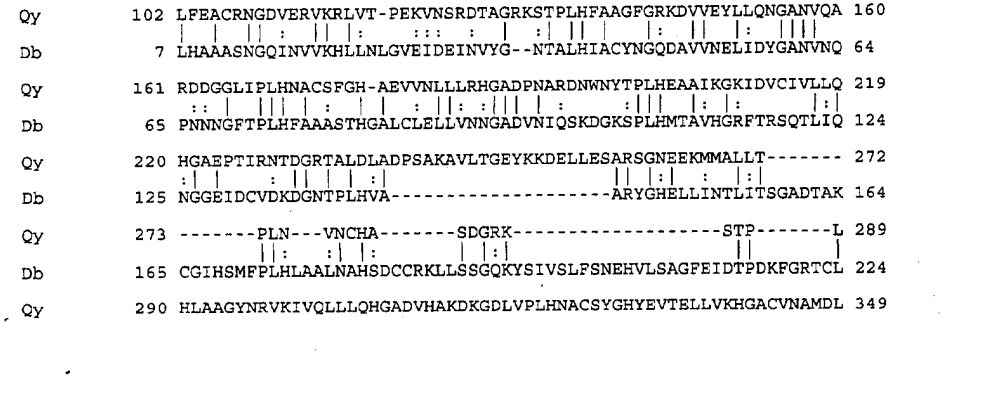
Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 836

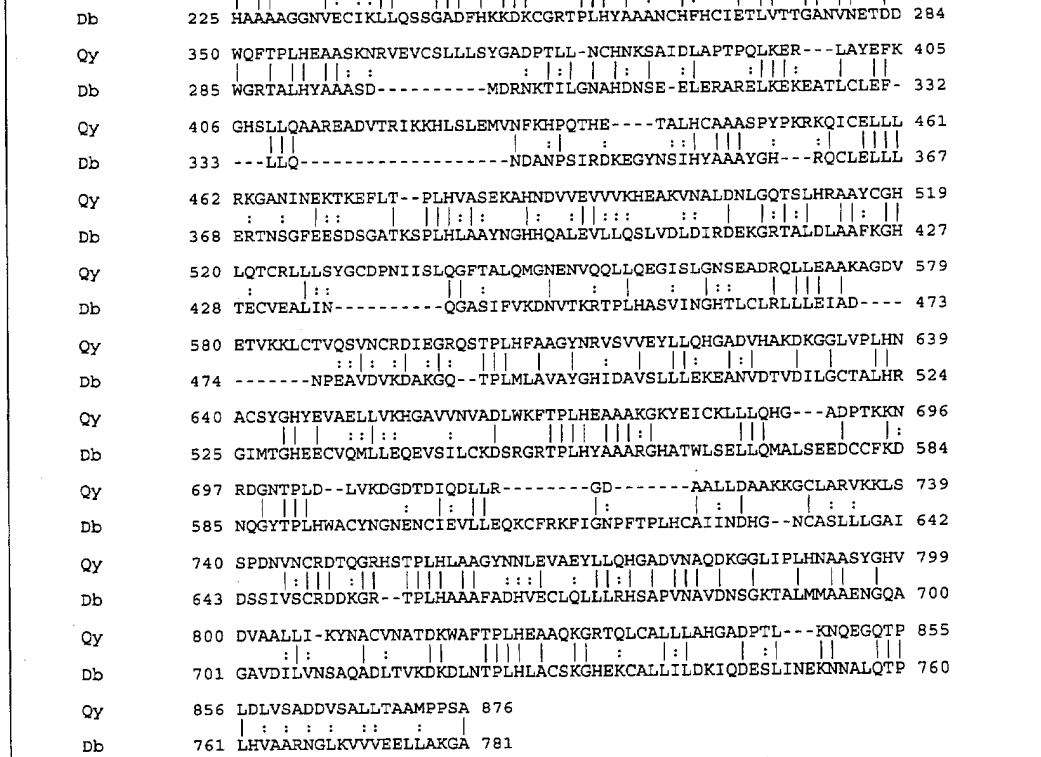
Db        614 ASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKEISG 673
```



RESULT 14
T42691
hypothetical protein DKFZp434D2328.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T42691
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A;Reference number: Z22230
A;Accession: T42691
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-791 <AAA>
A;Cross-references: EMBL:AL133087
A;Experimental source: adult testis; clone DKFZp434D2328
C;Genetics:
A;Note: DKFZp434D2328.1

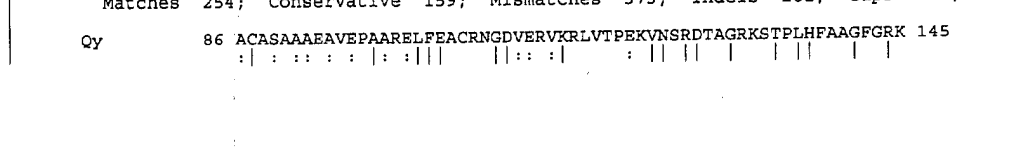
Query Match 10.2%; Score 657; DB 2; Length 791;
Best Local Similarity 28.2%; Pred. No. 2.2e-32;
Matches 243; Conservative 115; Mismatches 331; Indels 172; Gaps 29;

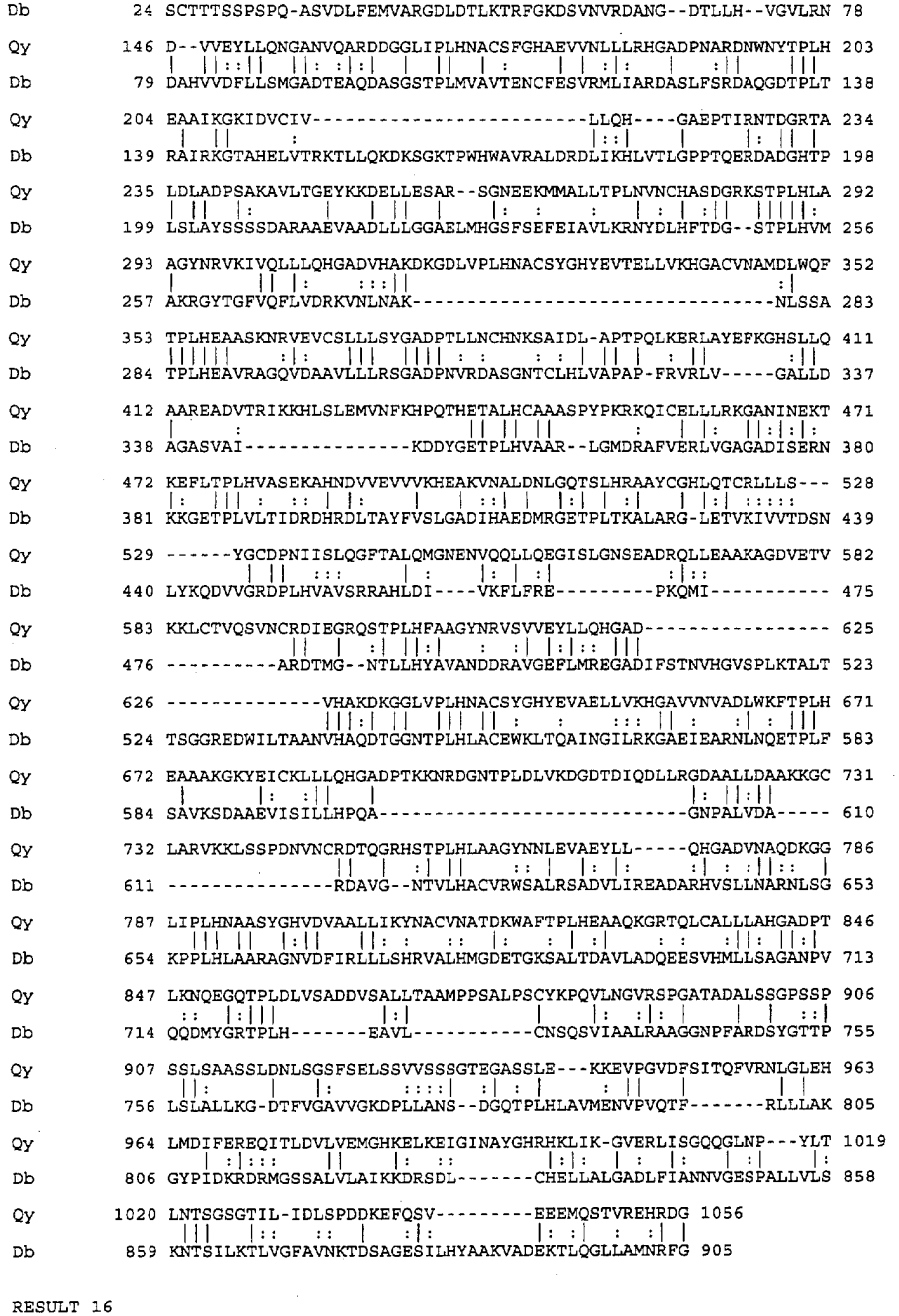




RESULT 15
H71274
probable ankyrin - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: H71274
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: H71274
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-934 <COL>
A;Cross-references: GB:AE001254; GB:AE000520; NID:g3323148; PIDN:AAC65803.1; PID:g332314
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0835
C;Superfamily: syphilis spirochete probable ankyrin; ankyrin repeat homology
F;281-313/Domain: ankyrin repeat homology <AN22>

Query Match 8.7%; Score 563.5; DB 1; Length 934;
Best Local Similarity 23.8%; Pred. No. 1.7e-26;
Matches 254; Conservative 159; Mismatches 373; Indels 281; Gaps 36;

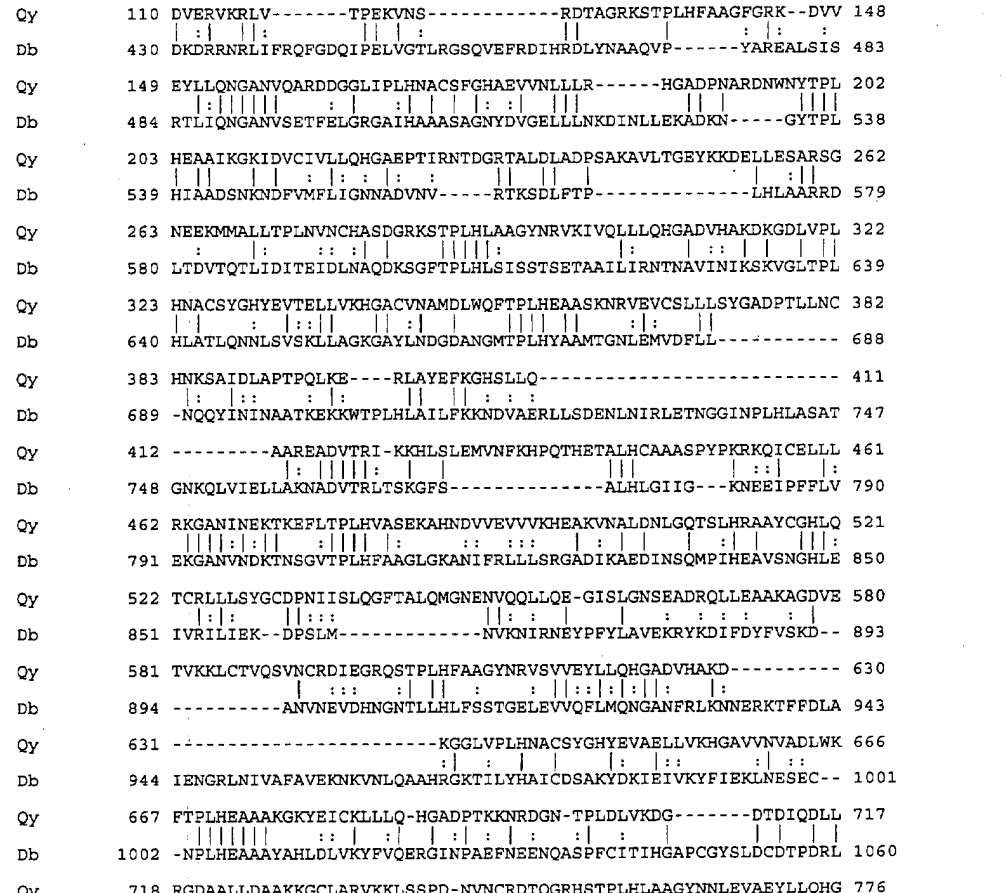




RESULT 16
S30355

alpha-latroinsectotoxin precursor - black widow spider (fragment)
C;Species: Latrodectus mactans tredecimguttatus (black widow spider)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Sep-1999
C;Accession: S30355; S29864
R;Kiyatkin, N.; Dulubova, I.; Grishin, E.
Eur. J. Biochem. 213, 121-127, 1993
A;Title: Cloning and structural analysis of alpha-latroinsectotoxin cDNA. Abundance of ar
A;Reference number: S30355; MUID:93238678; PMID:8477689
A;Accession: S30355
A;Molecule type: mRNA
A;Residues: 1-1411 <KIY>
A;Cross-references: EMBL:Z14086; NID:g9536; PIDN:CAA78464.1; PID:g9537
A;Note: it is uncertain whether Met-26 is the initiator or whether translation is initiat
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: presynaptic neurotoxin
F;533-565/Domain: ankyrin repeat homology <AN1>
F;1001-1032/Domain: ankyrin repeat homology <AN2>
F;1146-1178/Domain: ankyrin repeat homology <AN3>

Query Match 8.5%; Score 552; DB 2; Length 1411;
Best Local Similarity 23.4%; Pred. No. 1.6e-25;
Matches 251; Conservative 141; Mismatches 396; Indels 284; Gaps 37;



```
Db       1061 EVVEYLSD------------KIPDINGKC-DVQ--ENTPITVAIFANKVSILNYLVGIG 1104

Qy        777 ADVNAQDKGGLIPLHNAASYGHVDVAALLIK-YNACVNATDKWAFTPLHEAAQKGRTQLC 835

Db       1105 ADPNQQVDGD-PPLYIAARQGRFEIVRCLIEVHKVDINTRNKERFTALHAAARNDFMDVV 1163

Qy        836 ALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGAT 895

Db       1164 KYLVRQGADVNAKGIDDLRPID-IAGEKAKAYLQSS---RFLRSGHSFQ-SNEIDSFGNT 1218

Qy        896 ADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQF 955

Db       1219 IHGIS---------MSARTNDKLTQQISSKGTRSDSNSTEGKMHSENVHVRSIDVN---- 1265

Qy        956 VRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHK------------LIK 1002

Db       1266 ----GALLLLDFMIRVFAS-----------KKTNFAPYGSRIKTRSAAEAQAEALIMTE 1309

Qy       1003 GVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHR 1054

Db       1310 RFENLLSGLIG----------------DPIPDSIDFSNVHSKIYKAIMSGR 1344


RESULT 17
D82654
ankyrin-like protein XF1640 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82654
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82654
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1058 <SIM>
A;Cross-references: GB:AE003991; GB:AE003849; NID:g9106696; PIDN:AAF84449.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1640

  Query Match             7.9%;  Score 507.5;  DB 2;  Length 1058;
  Best Local Similarity   26.3%;  Pred. No. 5.8e-23;
  Matches  268;  Conservative  116;  Mismatches  395;  Indels  239;  Gaps   40;

Qy         47 GGLALLLAGPGLLLRLLALLLAVAAARIMSG-----RRCAGGGAAC------------A 88

Db        119 GSVAVALAWPGLLTLSLRWVLAVVAALLAPAMHVLLQRVQSRGLTARVEPPTSREPLPEV 178

Qy         89 SAAAEAVEP-----AARELFEACRNGDVERVKRLV----TPEKV---NSRDTAGRKSTPL 136

Db        179 TASPETVPPIPSDRLVADLYEAARTGRVDRALQLLEAGADPHALPFDNARD----QRSLM 234

Qy        137 HFAAGFGRKDVVEYLLQNGANVQARDDG--GLIPLHNACSFGHAEVVNLLLRHGADPNAR 194

Db        235 VLAAILPDLRLLLRELIVRRVDVNQRHRGMTALLAATRDSWHGRSEAVQTLLANGADPRVV 294
```

```
Qy        195 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 254

Db        295 DADRNTPLHHAARSSDPGVAALLRDAAAELDALNCDG------LSPLAAACQVGNWRLAK 348

Qy        255 -LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAG---YNRVKIVQLLLQHGAD 310

Db        349 FLLE--RGAQPE-------PIG----------GTPVLLAAAGTDEDDPTGVLLLLKHKAR 389

Qy        311 VHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLL 370

Db        390 VDARDRLRRTALHEAAEAGHSEIVKALLEAAANLEVRDMAGRTPLLEAARQARIVVLECL 449

Qy        371 LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMV 430

Db        450 LSHKADRLAVDSEGRNAVFLA---CLSERVT--------------LPLIRRLLEIGVPVV 492

Qy        431 NFKH---PQTHETALHC----AAASP-YPKRKQICE-------------LLLRKGANINE 469

Db        493 ADKHGHRPVDIAAKVGCWSIVSLLDPDYPLPVVVSESVSVGGVLDRPPLTLLRDG----- 547

Qy        470 KTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSY 529

Db        548 --------LQLGSFGSLSELIELC--------SAEELG-TLLHDPHFALNPDVLDWLLTH 590

Qy        530 GCDPNIISLQG----FTALQMGNENVQQL---LQEGISLGNSEADRQLLEAAKAGDVETV 582

Db        591 GAAPTVPNSSGNVPMFSLLARGIEAVPALKVFLRHGVS------------PAGAGGLGLW 638

Qy        583 KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACS 642

Db        639 LAACMQCDAASRSLE-------QFAC---------ELLEHGADPLMPSPDGDSPLVLTVR 682

Qy        643 YGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTP 702

Db        683 FGWLRLQQALLEAGVNCEVRDSYGMTALHQATVLGREAALKLLVMHGANPDARTPDGQTS 742

Qy        703 LDL-VKDGDTDIQDLL-----------RGDAALLDAAKKG---CLARVKKLSSPDNVNCR 747

Db        743 LGMALSSGRRDLATWLDWRSWPLPRRPLRQADVPDAAMRGDTEAVRRLLDLGLP--VDAV 800

Qy        748 DTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPL----------------- 790

Db        801 DAQG--CTALLHAAGGGHLQVVRLLLMRGADLQHAANNGATPLSAAVSMRQNEIIPALLD 858

Qy        791 ------HN----------AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKG---- 830

Db        859 AGAPLEHRFPEEMTVLMLAAALGFPDIVMRLIAAGADVHASNVQGFAALHCAALYGFSAR 918

Qy        831 -RTQLCAL---LLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 886

Db        919 DKTRFLALLDALLLAGANPGQLADGKITPLLL--------LLGARAERGAV--CDEQVVL 968

Qy        887 NGVR---SPGATADALSSGPSSPSSLSAASSLDNL------SGSFSELSSVVSSSGTE 935

Db        969 AGVERLLEEGDTLDVVDTRGFGPLHLAALHGLPLLVKCLLRAGADSERRDALNRSPRE 1026


RESULT 18
T30255
inversin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C;Accession: T30255
R;Morgan, D.; Turnpenny, L.; Goodship, J.; Dai, W.; Majumder, K.; Matthews, L.; Gardner,
Nature Genet. 20, 149-156, 1998
A;Title: Inversin, a novel gene in the vertebrate left-right axis pathway, is partially
A;Reference number: Z20793; MUID:98442650; PMID:9771707
A;Accession: T30255
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1062 <MOR>
A;Cross-references: EMBL:AJ010902; NID:g3757671; PIDN:CAA09388.1; PID:g3757672
A;Experimental source: strain FBV/N
```

C;Genetics:
A;Gene: invs
A;Map position: 4

Query Match 7.7%; Score 494.5; DB 2; Length 1062;
Best Local Similarity 24.0%; Pred. No. 3.7e-22;
Matches 210; Conservative 101; Mismatches 290; Indels 275; Gaps 23;

```
Qy        177 HAEVVN--------LLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 228
              ||  :|        |:: : |  :  | :  |||    :  ::|    ||: ||:    :
Db         19 HAAAINGDKGALQKLIVGNSALRDKEDRFGRTPLMYCVLADRVDCADALLKAGADVNKTD 78

Qy        229 TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP 288
                ||||  ||                    |: ||   |  |||    |    |  : ||
Db         79 HSRRTALHLA-------------------AQKGNYRFMKLLLT-RRANWMQKDLEEMTP 117

Qy        289 LHLAAGYNRVKIVQLLLQHGA--DVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNA 346
              |||:  :   | : |||:  |  :|  :||      || :  | : |  :||:|| : :
Db        118 LHLSTRHRSPKCLALLLKFMAPGEVDTQDKNKQTALHWSAYYNNPEHAKLLIKHDSNIGI 177

Qy        347 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKG 406
              |:    ||| ||:                    ||: ::   :  :| |||   |
Db        178 PDVEGKIPLHWAANHK------------DPSAVHT-VRCILDAAPTESL---------- 213

Qy        407 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 466

Db        214 ------------------------------------------------------------ 213

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEA-KVNALDNLGQTSLHRAAYCGHLQTCRL 525
              :| :  |  |||| |    :  ||:|:  :|:  : : ||| :| || ||  || |   |
Db        214 LNWQDYEGRTPLHFAVADGNLTVVDVLTSYESCNITSYDNLFRTPLHWAALLGHAQIVHL 273

Qy        526 LLSYGCDPNIIS-LQGFTALQMG-----NENVQQLLQEGISLGNS--EADRQLLEAAKAG 577
              ||          | |  || | |              | ||:  ||      :|  |      : ||  |
Db        274 LLERNKSGTIPSDSQGATPLHYAAQSNFAETVKVFLQHPSVKDDSDQEGRTSFMWAAGKG 333

Qy        578 DVETVKKLCTVQS---VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGL 634
              : : :: : :::|   :|  |  |    | || ||    || |: || : | | | |
Db        334 NDDVLRTMLSLKSDIDINMSDKYG--GTALHAAALSGHVSTVKLLLDNDAQVDATDVMKH 391

Qy        635 VPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTK 694
              ||  ||  || :| : |:| || |:: |    : || ||  |  ::|::|:::  :|
Db        392 TPLFRACEMGHRDVIQTLIKGGARVDLVDQDGHSLLHWAALGGNADVCQILIENKINP-- 449

Qy        695 KNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
                                                          | :|  ||
Db        450 ------------------------------------------------NVQDYAGR-- 457

Qy        755 TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVN- 813
              ||:  ||   |:     |::: || | ||| |   || : : |::|   ||: : |  |
Db        458 TPIQSAAYGGNINCMAVLMENNADPNIQDKEGRTALHWSCNNGYLDAIKLLLDFAAFPNQ 517

Qy        814 -ATDKWAFTPLHEAAQKGRTQLCALLLAHGA---------------------------- 843
               ::  :|||  |   | ::   :| |||
Db        518 MENNEERYTPLDYALLGERHEVIQFMLEHGALSIAAIQDIAAFKIQAVYKGYKVRKAFRD 577

Qy        844 -------------DPTLK-----------NQEGQTPLDLVSADDVSALLTAAMPPS--- 875
                          |  |            |:||   |   :  |:      |||
Db        578 RKNLLMKHEQLRKDAAAKKREEENKRKEAEQQKGQLDTDPPRSHCSSSAPVLPCPPSPQN 637

Qy        876 ------ALPSCYKPQVLNGVRSPGATADALSSGPSS--PSSLSAASSLDNLSGSFSELSS 927
                   | ||  :|   : |:||   |  | |   |  |  |  :| |: |   |
Db        638 EASKQDATPS-KQPPASHTVQSP----DPEHSRPPGRCPGRASQGDSSIDLQGTASRKPS 692

Qy        928 VV-------------------SSSGTEGASSLEKK 943
                                  :||  :| ||:||:
Db        693 ETPIEHCRGPSACVHPRSWEGGNSSKNQGTSSVEKR 728
```

RESULT 19
T14151
Inv protein - mouse
C;Species: Mus musculus (house mouse)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14151
R;Yokoyama, T.; Mochizuki, T.
submitted to the EMBL Data Library, November 1997
A;Reference number: Z17888
A;Accession: T14151
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1062 <YOK>
A;Cross-references: EMBL:AF034860; NID:g3560522; PID:g3560523; PIDN:AAC34976.1
C;Genetics:
A;Gene: inv
A;Map position: 4

Query Match 7.6%; Score 489.5; DB 2; Length 1062;
Best Local Similarity 24.2%; Pred. No. 7.4e-22;
Matches 212; Conservative 98; Mismatches 291; Indels 275; Gaps 24;

```
Qy        177 HAEVVN--------LLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 228
              ||  ||        |:: : |  :  | :  |||    :  ::|    ||: ||:    :
Db         19 HAAAVNGDKGALQRLIVGNSALRDKEDRFGRTPLMYCVLADRVDCADALLKAGADVNKTD 78

Qy        229 TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP 288
                ||||  ||                    |: ||   |  |||    |    |  : ||
Db         79 HSRRTALHLA-------------------AQKGNYRFMKLLLT-RRANWMQKDLEEMTP 117

Qy        289 LHLAAGYNRVKIVQLLLQHGA--DVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNA 346
              |||:  :   | : |||:  |  :|  :||      || :  | : |  :||:|| : :
Db        118 LHLSTRHRSPKCLALLLKFMAPGEVDTQDKNKQTALHWSAYYNNPEHAKLLIKHDSNIGI 177

Qy        347 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKG 406
              |:    ||| ||:                    ||: ::   :  :| |||   |
Db        178 PDVEGKIPLHWAANHK------------DPSAVHT-VRCILDAAPTESL---------- 213

Qy        407 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 466

Db        214 ------------------------------------------------------------ 213

Qy        467 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEA-KVNALDNLGQTSLHRAAYCGHLQTCRL 525
              :| :  |  |||| |    :  ||:|:  :|:  : : ||| :| || ||  || |   |
Db        214 LNWQDYEGRTPLHFAVADGNLTVVDVLTSYESCNITSYDNLFRTPLHWAALLGHAQIVHL 273

Qy        526 LLSYGCDPNIIS-LQGFTALQMG-----NENVQQLLQEGISLGNS--EADRQLLEAAKAG 577
              ||          | |  || | |              | ||:  ||      :|  |      : ||  |
Db        274 LLERNKSGTIPSDSQGATPLHYAAQSNFAETVKVFLQHPSVKDDSDLEGRTSFMWAAGKG 333

Qy        578 DVETVKKLCTVQS---VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGL 634
              : : :: : :::|   :|  |  |    | || ||    || |: || : | | | |
Db        334 NDDVLRTMLSLKSDIDINMSDKYG--GTALHAAALSGHVSTVKLLLDNDAQVDATDVMKH 391

Qy        635 VPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTK 694
              ||  ||  || :| : |:| || |:: |    : || ||  |  ::|::|:::  :|
Db        392 TPLFRACEMGHRDVIQTLIKGGARVDLVDQDGHSLLHWAALGGNADVCQILIENKINP-- 449

Qy        695 KNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS 754
                                                          | :|  ||
Db        450 ------------------------------------------------NVQDYAGR-- 457

Qy        755 TPLHLAA--GYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 812
              |||  ||  || |    |::: || | ||| |   || : : |::|   ||: : |
Db        458 TPLQCAAYGGYINCMAV--LMENNADPNIQDKEGRTALHWSCNNGYLDAIKLLLDFAAFP 515

Qy        813 N--ATDKWAFTPLHEAAQKGRTQLCALLLAHGA---------------------------- 843
              |    ::  :|||  |    | ::   :| |||
Db        516 NQMENNEERYTPLDYALLGERHEVIQFMLEHGALSIAAIQDIAAFKIQAVYKGYKVRKAF 575

Qy        844 --------------DPTLK-----------NQEGQTPLDLVSADDVSALLTAAMPPS- 875
                           |  |            |:||   |   :  |:      |||
Db        576 RDRKNLLMKHEQLRKDAAAKKREEENKRKEAEQQKGQLDTDPPRSHCSSSAPVLPCPPSP 635

Qy        876 --------ALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSS 927
                     | ||  :|   : |:||       |       |  |   |  :| |: |   |
Db        636 QNEGSKQDATPS-KQPPASHTVQSPDPEHSRLPG--RCPGRASQGDSSIDLQGTASRKPS 692
```

```
Qy        928 VV-------------------SSSGTEGASSLEKK 943
                                   :||  :| ||:||:
Db        693 ETPIEHCRGPSACVHPRSWEGGNSSKNQGTSSVEKR 728
```

RESULT 20
S11527
alpha-latrotoxin precursor - black widow spider
C;Species: Latrodectus mactans tredecimguttatus (black widow spider)
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Sep-1999
C;Accession: S11527; JN0313
R;Kiyatkin, N.I.; Dulubova, I.E.; Chekhovskaya, I.A.; Grishin, E.V.
FEBS Lett. 270, 127-131, 1990
A;Title: Cloning and structure of cDNA encoding alpha-latrotoxin from black widow spider
A;Reference number: S11527; MUID:91031994; PMID:1977615
A;Accession: S11527
A;Molecule type: mRNA
A;Residues: 1-1401 <KIY>
A;Cross-references: EMBL:X55009; NID:g9568; PIDN:CAA38753.1; PID:g9569
A;Note: it is uncertain whether Met-1, Met-8 or Met-18 is the initiator
R;Volkova, T.M.; Galkina, T.G.; Kudelin, A.B.; Grishin, E.V.
Bioorg. Khim. 17, 437-441, 1991
A;Title: Tryptic fragments structure of the black widow spider venom neurotoxin.
A;Reference number: JN0313; MUID:91362695; PMID:1888339
A;Note: English title and abstract; article in Russian
A;Accession: JN0313
A;Molecule type: protein
A;Residues: 21-33,'X',35-40,'X',43-50;82-90,'X',92-99;107-113;114-126;146-151;171-178;19
84-792;881-891;906-913;914,1015-1021;1062-1065,'X',1067-1075;1080-1084;1087-1096;1183-11
A;Note: the sequence of the amino terminal peptide has been corrected with tryptic pepti
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: neurotoxin; venom
F;1-20/Domain: propeptide #status predicted <PRO>
F;21-1401/Product: alpha-latrotoxin #status experimental <MAT>
F;525-557/Domain: ankyrin repeat homology <AN7>
F;862-894/Domain: ankyrin repeat homology <AN3>
F;1005-1036/Domain: ankyrin repeat homology <AN21>

  Query Match   7.5%; Score 485; DB 2; Length 1401;
  Best Local Similarity   23.5%; Pred. No. 2.2e-21;
  Matches 191; Conservative 100; Mismatches 302; Indels 220; Gaps   22;

```
Qy        100 RELFEACRNGD-----VERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYL--L 152
Db        456 RDLYNAASNPDSAVGFKEFTK--LNYDGANIRATFDHGRTVFHAAAKSGNDKIMFGLTFL 513

Qy        153 QNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGKI 211
Db        514 AKSTELNQPDKKGYTPIHVAADSGNAGIVNLLIQRGVSINSKTYHFLQTPLHLAAQRGFV 573

Qy        212 DVCIVLLQHGAEPTI----RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKM 267
Db        574 TTFQRLME---SPEININERDKDGFTP--------------------LHYAIRGGERIL 609

Qy        268 MALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAKDKGDLVPLHNAC 326
Db        610 EAFLNQISIDVNAKSNTGLTPFHLAIIKNDWPVASTLLGSKKVDINAVDENNITALHYAA 669

Qy        327 SYGHYEVTE------------------------LLVKH----------GACVNAMDLW 350
Db        670 ILGYLETTKQLINLKEINANVVSSPGLLSALHYAILYKHDDVASFLMRSSNVNVNLKALG 729

Qy        351 QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL 410
Db        730 GITPLHLAVIQGRKQILSLMFDIG---------------------------------- 753

Qy        411 QAAREADVTRIKKHLSLEMVNFKHPQTHE--TALHCAAASPYPKRKQICELLLRKGANIN 468
Db        754 ------------------VNIEQ-KTDEKYTPLHLAAMSKYP---ELIQILLDQGSNFE 790

Qy        469 EKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS 528
```

```
Db        791 AKTNSGATPLHLATFKGKSQAALILLNNEVNWRDTDENGQMPIHGAAMTGLLDVAQAIIS 850

Qy        529 YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 588
Db        851 --IDATVV---------------------------------------------------- 856

Qy        589 QSVNCRDIEGRQS-TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 647
Db        857 ------DIEDKNSDTPLNLAAQNSHIDVIKYFIDQGADINTRNKKGLAPLLAFSKKGNLD 910

Qy        648 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 707
Db        911 MVKYLFDKNANVYIADNDGMNFFYYAVQNGHLNIVKYAMSEKDKFEWSNTDNN------R 964

Qy        708 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHST--PLHLAAGYNN 765
Db        965 RDECPNEECAISHFAVCDAVQFDRIEIVKYFVG--------TLGNFAICGPLHQAARYGH 1016

Qy        766 LEVAEYLLQH---GADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTP 822
Db       1017 LDIVKYLVEEEFLSVDGSKTD----TPLCYASENGHFTVVQYLVSNGAKVNHDCGNGMTA 1072

Qy        823 LHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 855
Db       1073 IDKAITKNHLQVVQFLAANGVDFRRKNSRGTTP 1105
```

RESULT 21
T19552
hypothetical protein C29E6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19552
R;Dobson, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z19141
A;Accession: T19552
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1188 <WIL>
A;Cross-references: EMBL:Z72504; PIDN:CAA96603.1; GSPDB:GN00022; CESP:C29E6.2
A;Experimental source: clone C29E6
C;Genetics:
A;Gene: CESP:C29E6.2
A;Map position: 4
A;Introns: 147/2; 473/3; 500/1; 529/1; 584/2; 688/2; 839/3; 975/3; 1132/3

  Query Match   7.2%; Score 463; DB 2; Length 1188;
  Best Local Similarity   22.8%; Pred. No. 3.8e-20;
  Matches 210; Conservative 137; Mismatches 384; Indels 192; Gaps   26;

```
Qy        107 RNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVV-EYLLQNGANVQARDDGG 165
Db         13 RQASIFRVAEL------DARTEADNLRSIIHQSAREGNVNALQEALLKAPLAVNAQDGDF 66

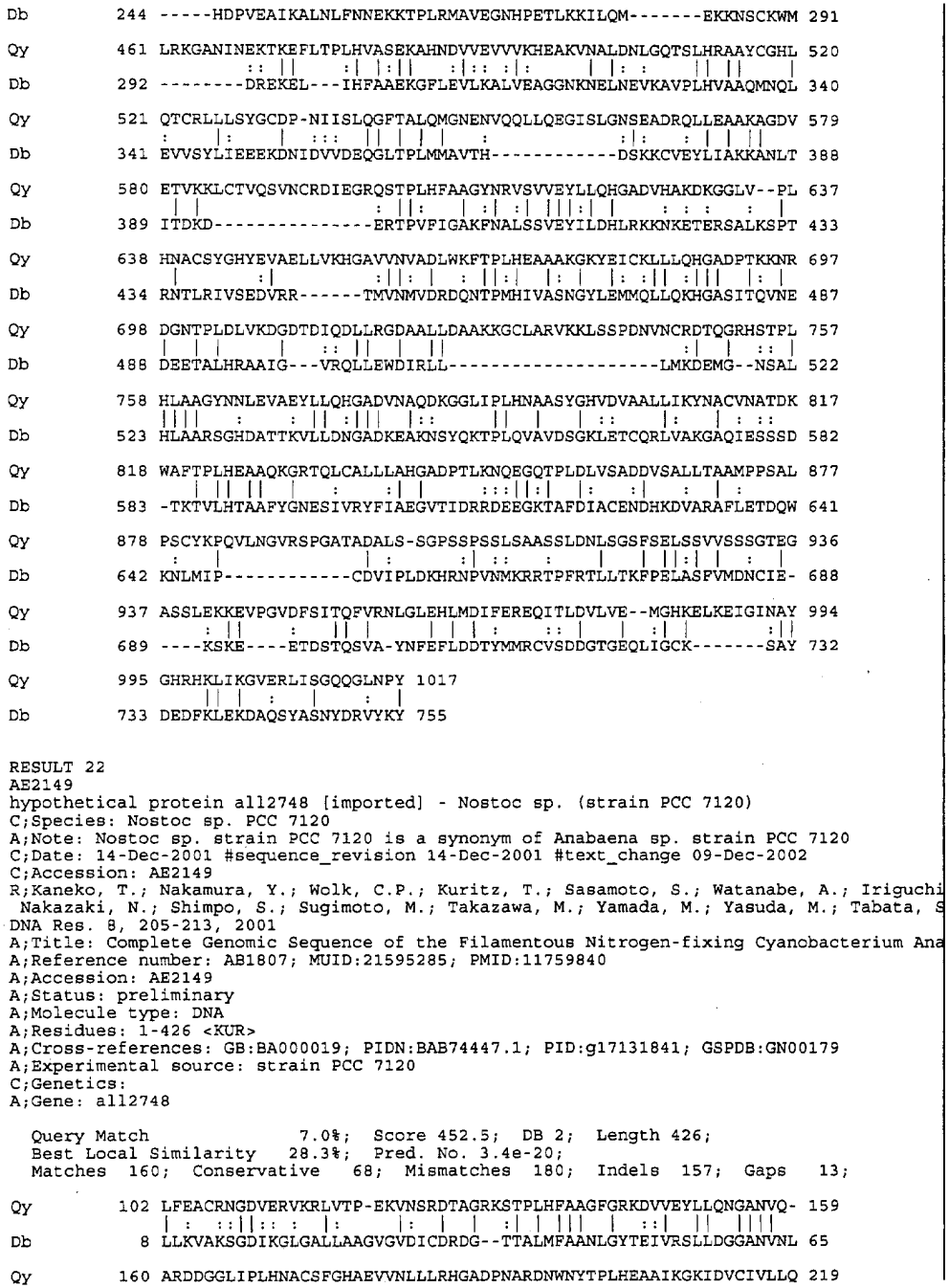
Qy        166 LIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAA--IKGKIDVCIVLLQHGAE 223
Db         67 MTPLHYAARYGNYDAVKLLLSKNALPNTKNREGDTPLHIASKYIYGYSDICSII------ 120

Qy        224 PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGN--EEKMMALLTPLNVNCHAS 281
Db        121 ---------------------------DEDQADSARKYNTATKKIINALVSENAEIDPV 152

Qy        282 DGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHG 341
Db        153 NKYQLTPLHYAAMKSNFSALHALIKLKADVDAEDDNKMTPLLLACVHGSQEIIQELIKAN 212

Qy        342 ACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLA 401
Db        213 SNVTKRDQRLNTVFHIVALRGEPEYLEMMMD---------------------------- 243

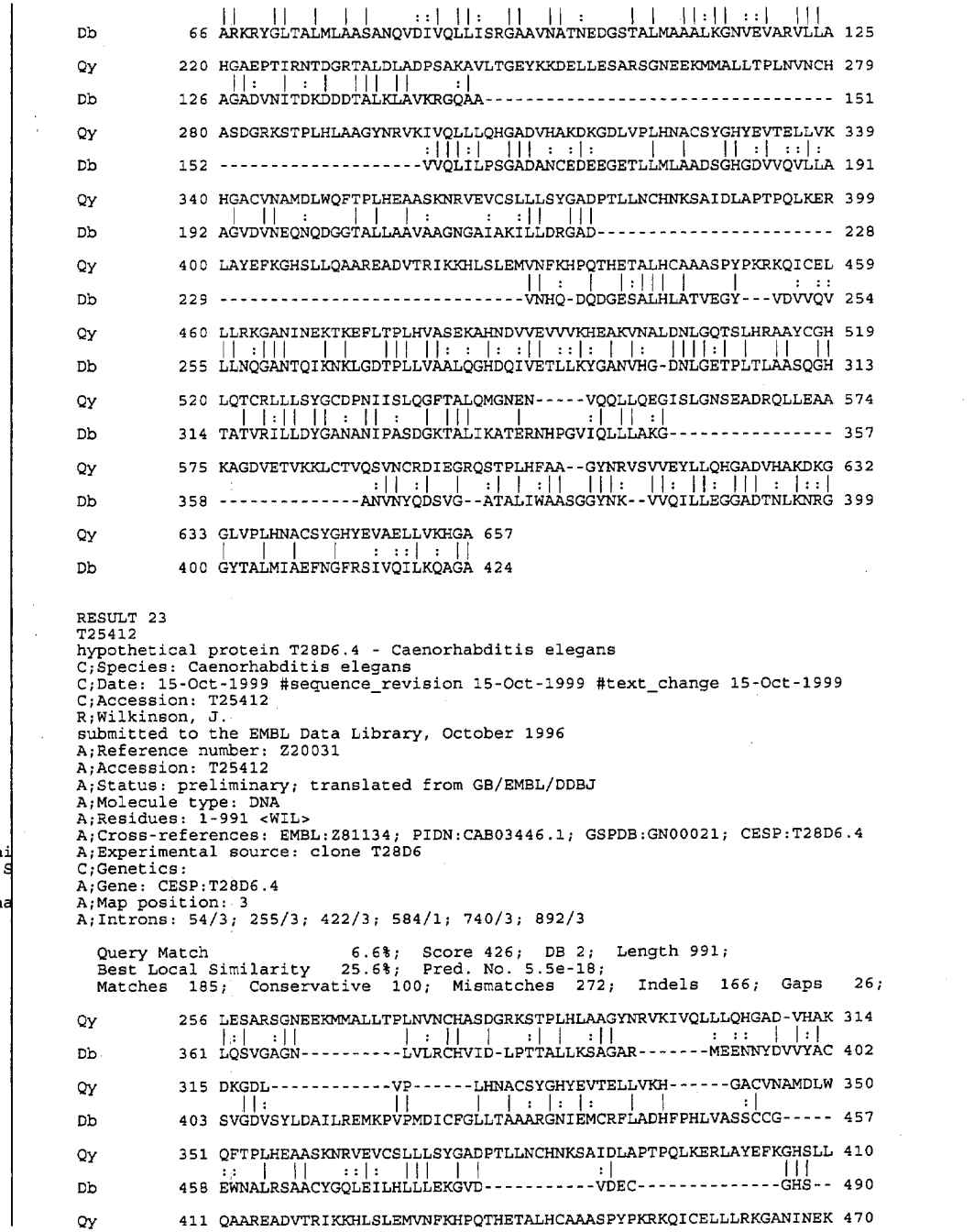
Qy        402 YEFKGHSLLQAAREADVTRIKKHLSLEM-VNFKHPQTHETALHCAAASPYPKRKQICELL 460
```



RESULT 22
AE2149
hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2149
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-426 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74447.1; PID:g17131841; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2748

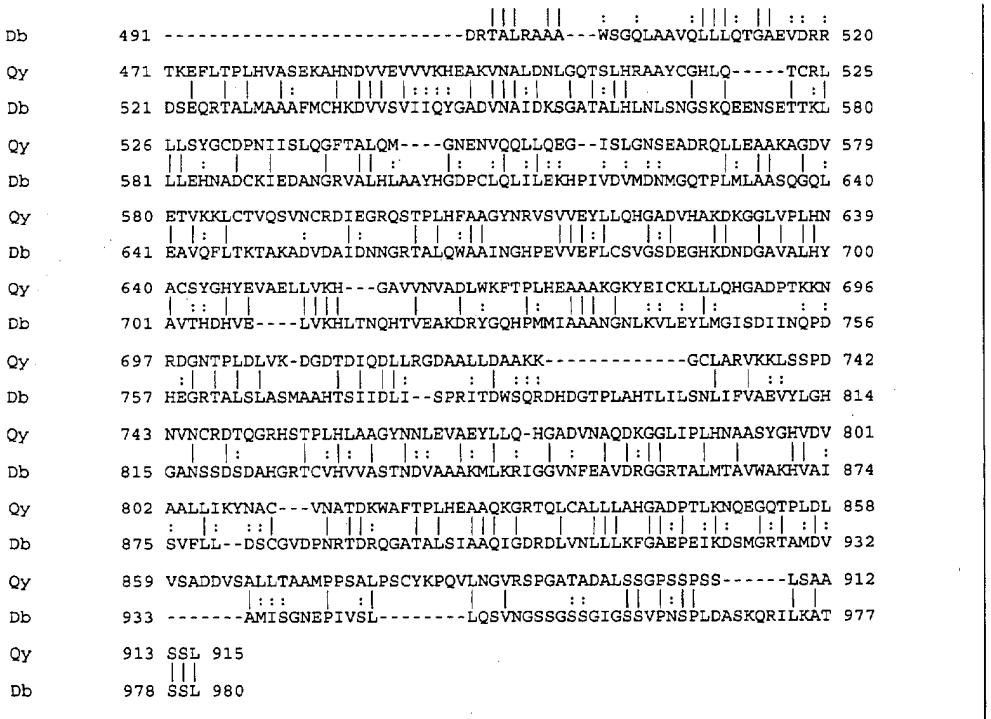
Query Match 7.0%; Score 452.5; DB 2; Length 426;
Best Local Similarity 28.3%; Pred. No. 3.4e-20;
Matches 160; Conservative 68; Mismatches 180; Indels 157; Gaps 13;





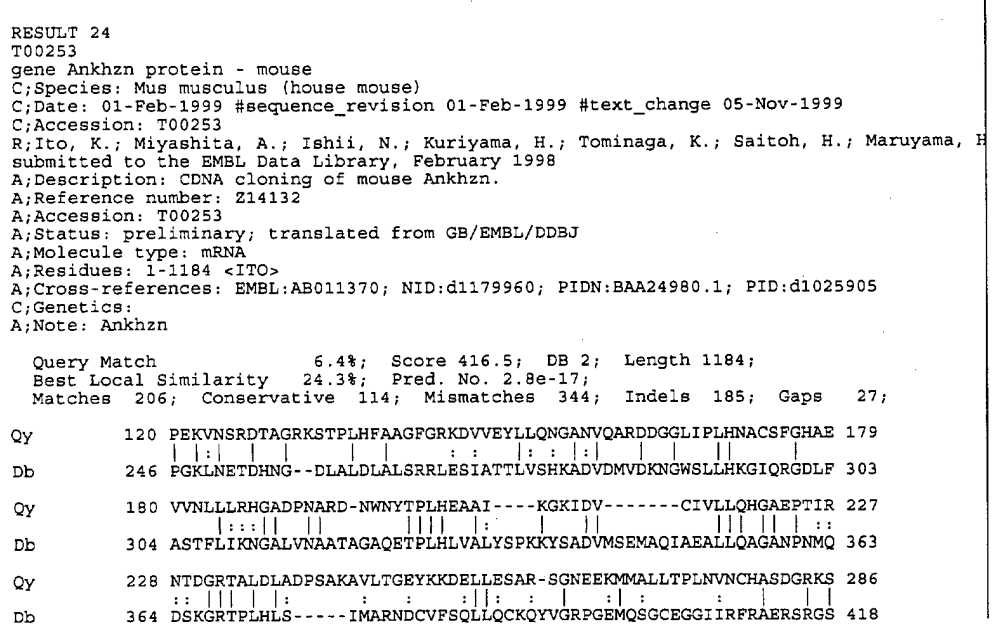
RESULT 23
T25412
hypothetical protein T28D6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25412
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20031
A;Accession: T25412
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-991 <WIL>
A;Cross-references: EMBL:Z81134; PIDN:CAB03446.1; GSPDB:GN00021; CESP:T28D6.4
A;Experimental source: clone T28D6
C;Genetics:
A;Gene: CESP:T28D6.4
A;Map position: 3
A;Introns: 54/3; 255/3; 422/3; 584/1; 740/3; 892/3

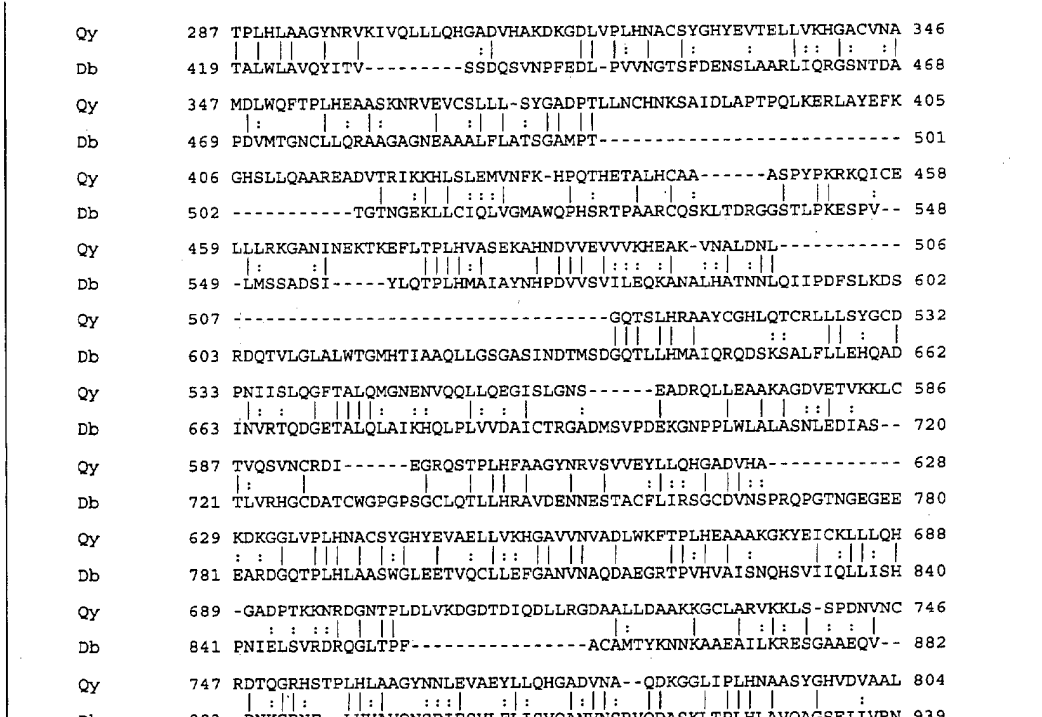
Query Match 6.6%; Score 426; DB 2; Length 991;
Best Local Similarity 25.6%; Pred. No. 5.5e-18;
Matches 185; Conservative 100; Mismatches 272; Indels 166; Gaps 26;



RESULT 24
T00253
gene Ankhzn protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 05-Nov-1999
C;Accession: T00253
R;Ito, K.; Miyashita, A.; Ishii, N.; Kuriyama, H.; Tominaga, K.; Saitoh, H.; Maruyama, H
submitted to the EMBL Data Library, February 1998
A;Description: CDNA cloning of mouse Ankhzn.
A;Reference number: Z14132
A;Accession: T00253
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1184 <ITO>
A;Cross-references: EMBL:AB011370; NID:d1179960; PIDN:BAA24980.1; PID:d1025905
C;Genetics:
A;Note: Ankhzn

Query Match 6.4%; Score 416.5; DB 2; Length 1184;
Best Local Similarity 24.3%; Pred. No. 2.8e-17;
Matches 206; Conservative 114; Mismatches 344; Indels 185; Gaps 27;





RESULT 25
T24158
hypothetical protein R11A8.7b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T24158; T24179
R;Bardill, S.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19846
A;Accession: T24158
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2584 <WIL>
A;Cross-references: EMBL:Z70686; PIDN:CAB54289.1; GSPDB:GN00022; CESP:R11A8.7b
A;Experimental source: clone R10H10
R;Cummings, P.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19849
A;Accession: T24179
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2584 <WI2>
A;Cross-references: EMBL:Z70310; PIDN:CAB54294.1; GSPDB:GN00022; CESP:R11A8.7b

```
A;Experimental source: clone R11A8
C;Genetics:
A;Gene: CESP:R11A8.7b
A;Map position: 4
A;Introns: 24/3; 80/3; 137/2; 177/1; 218/3; 348/3; 507/2; 709/3; 753/2; 924/1; 1020/3; 1

  Query Match             6.2%;  Score 399;  DB 2;  Length 2584;
  Best Local Similarity   20.8%;  Pred. No. 1.1e-15;
  Matches  251;  Conservative  179;  Mismatches  477;  Indels  300;  Gaps   39;

Qy          9 AGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLA 68
              || || | |              :      |  :|    |  | ||     : :: :|:
Db        466 AGDQGDQAGV----------VKLLLSKHAEVDVSNPDTGDTPLSLAARNGYIAIMKMLI- 514

Qy         69 VAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDT 128
                        ||   |  :  |          || ||| :| :: ::    |   :|
Db        515 ------------EKGGDLTAGKTSPIV--------EAARNGHLECIQFILAHCKTIPQDQ 554

Qy        129 AGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHG 188
                |    |  || ||   :|| ::: ||::    |      |  |    | ||| |||  |
Db        555 LSR---ALVSAADFGSLLIVEEVIRAGADLNFEQD-ERTALMKAAKGDHFEVVQLLLSKG 610

Qy        189 ADPNARDNWN-YTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG---------RTALDLA 238
              |  | : :  |  | |   :| :::    |:::||:| ::  ||           ::||
Db        611 ASVNFKSSKNDATALSLACSEGNMEIAEFLIRNGADPMLKMDDGVNCFMEVARHGSIDLM 670

Qy        239 D-----------------------------------PSAKAVLT---GEYKKDE----- 254
                                                  || : :|:   | | | :
Db        671 SLLVEFTKGNMPMDKDPPKLGITRCSSKNGKKRRKGMPSGQDMLSMFNGMYPKRKGSKQM 730

Qy        255 -LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQ--HGADV 311
               |  |  |  |  |:  |  :     | :  |||    |    :| |   ::  :|
Db        731 GLHEMPFSTQEIDMLTHLLKMQQQMVAYETHKSTETETA----EIKKVLRAIESVYGFTS 786

Qy        312 HAK-------DKGDLVPLHNACSYGHYEVTELLVKHG------ACVNAMDLWQFTPLHEA 358
                |        :: |:  |:|     : ::    || ||            ::|  |   :|
Db        787 EGKINFPPPPNRQDMDKLYNGELVPNIKLWAELVSHGWLEMERKIGRPIELSSFQQCNEG 846

Qy        359 ASKNRVEVCSL-----------------------------------LLSYGADPTLLNC 382
              |  |      |                                   | :  |  | ::
Db        847 HSTNAAAAVSAVAAAATGMDSQTYLASVFAKMNNGEEMPRVPATVGSLNAASAAMTGISF 906

Qy        383 HNKSAIDLAPTPQLKERLAYEFKG----------HSLLQAAREADVTRIKK--------H 424
              |:  |: |            :|  : |           | : : |  | : ::
Db        907 HSDDAMRLFGGASFATKLLSDNKKPCNHQQYASIHHIQEGAFRAALIKMSSMFRERNGCA 966

Qy        425 LSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKE---PLTPLHVA 481
              :|:  :    |  |:      : | |  :    |      |: :: | :    :    :
Db        967 ISVRDMESNFPIEHQEE-RFGPSKPIPSGPKKTSLTAPNPADTSDVTTKQPGAMKKDGLE 1025

Qy        482 SEKAHNDVVEVVVKHEA--KVNALD---NLGQTSLHRAA-----YCGHLQ------TCRL 525
              ::| :  ::::  : |   : |  :    :|  |: : |:      :|  ||      :
Db       1026 AKKIYPGIIKLAAEMEKLFRANPTESNRDLALTTAYIASALPDHFCSELQLESGDRILKK 1085

Qy        526 LLSYGCDP---NIISL--------QGFTALQMGNENV--QQLLQEGISLGNSEAD----- 567
              |||   :     :||          | : |:   :|:  ::: :: : |      |
Db       1086 LLSGLTEKQKNTVISRMRSVVNKESGSSLLRRSVDNLTDKRIKEDYLKLFRDSTDCAFYD 1145

Qy        568 ---------RQLLEAAKAGD------VETVKKLCTVQSVNCRDIEGR------------- 599
                      : :|  | |        : |  |    :||  :  :|:
Db       1146 KCVQEKNHLLKAIEIQKKGKTSSGTLISTSSKSLMAKSVQSQQQQGQLRRTHSEGDGAER 1205

Qy        600 ---------------QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 644
                            |||  |        :|| ||:  ||:   :|| | |  ||  | : |
Db       1206 AKSRSNAIDKATETTLETPLTIACANGHKDIVELLLKEGANIEHRDKKGFSPLIIAATAG 1265

Qy        645 HYEVAELLVK-HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 703
              |  | |:|:| | |:   :| | | |  | : |: :: :||| |||:   :|    |||
Db       1266 HSSVVEVLLKNHAAIEAQSDRTKDTALSLACSGGRKDVVELLLAHGANKEHRNVSDYTPL 1325

Qy        704 DLVKDGD-TDIQDLL--------------RGDAALLDAAKKGCLARVK-KLSSPDNVNCR 747
              |   |   :| ::|               | : |: |:  |      :  |    ::| :
Db       1326 SLASSGGYIEIVNMLLTAGSEINSRTGSKLGISPLMLASMNGHREATRVLLEKGSDINAQ 1385

Qy        748 DTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 807
                  | :| | ||:      || : || : |:|  : | || ||    || |:|||  |||
Db       1386 IETNR-NTALTLASFQGRTEVVKLLLAYNANVEHRAKTGLTPLMECASGGYVDVGNLLIA 1444

Qy        808 YNACVNAT--DKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVS 865
                |  ||:   :   | |  :|:||  :   :||   |   ::|::| | | |
Db       1445 AGADTNASPVQQTKDTALTISAEKGHEKFVRMLLNGDAAVDVRNKKGCTALWLACNGGYL 1504

Qy        866 ALLTAAMPPSALPSCYKPQVLNGVRSPGATA------------DALSSGPSSPSSLSA- 911
                :  | :  | | :     |  ||  |                ::    |:    : |
Db       1505 STAQALLEKGADPDMFD----NRKISPMMAAFRKGHVEIVKYMVNSAKQFPNEQDLIRAQ 1560

Qy        912 -ASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFER 970
               :  |::     |   ::  |:    | | |                 |  : |::: :
Db       1561 QTAETDDIKKKCGECIDIIRSAKKAQAESAE----------------LAAQKLLELIDE 1603

Qy        971 EQITLDV 977
                |::  :|
Db       1604 EKVQKEV 1610


RESULT 26
T24157
hypothetical protein R11A8.7a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T24157; T24177
R;Bardill, S.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19846
A;Accession: T24157
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2606 <WIL>
A;Cross-references: EMBL:Z70686; PIDN:CAA94615.1; GSPDB:GN00022; CESP:R11A8.7a
A;Experimental source: clone R10H10
R;Cummings, P.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19849
A;Accession: T24177
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2606 <WI2>
A;Cross-references: EMBL:Z70310; PIDN:CAA94370.1; GSPDB:GN00022; CESP:R11A8.7a
A;Experimental source: clone R11A8
C;Genetics:
A;Gene: CESP:R11A8.7a
A;Map position: 4
A;Introns: 24/3; 80/3; 137/2; 177/1; 218/3; 348/3; 507/2; 709/3; 753/2; 924/1; 1020/3; 1(

  Query Match             6.2%;  Score 399;  DB 2;  Length 2606;
  Best Local Similarity   20.8%;  Pred. No. 1.1e-15;
  Matches  251;  Conservative  179;  Mismatches  477;  Indels  300;  Gaps   39;

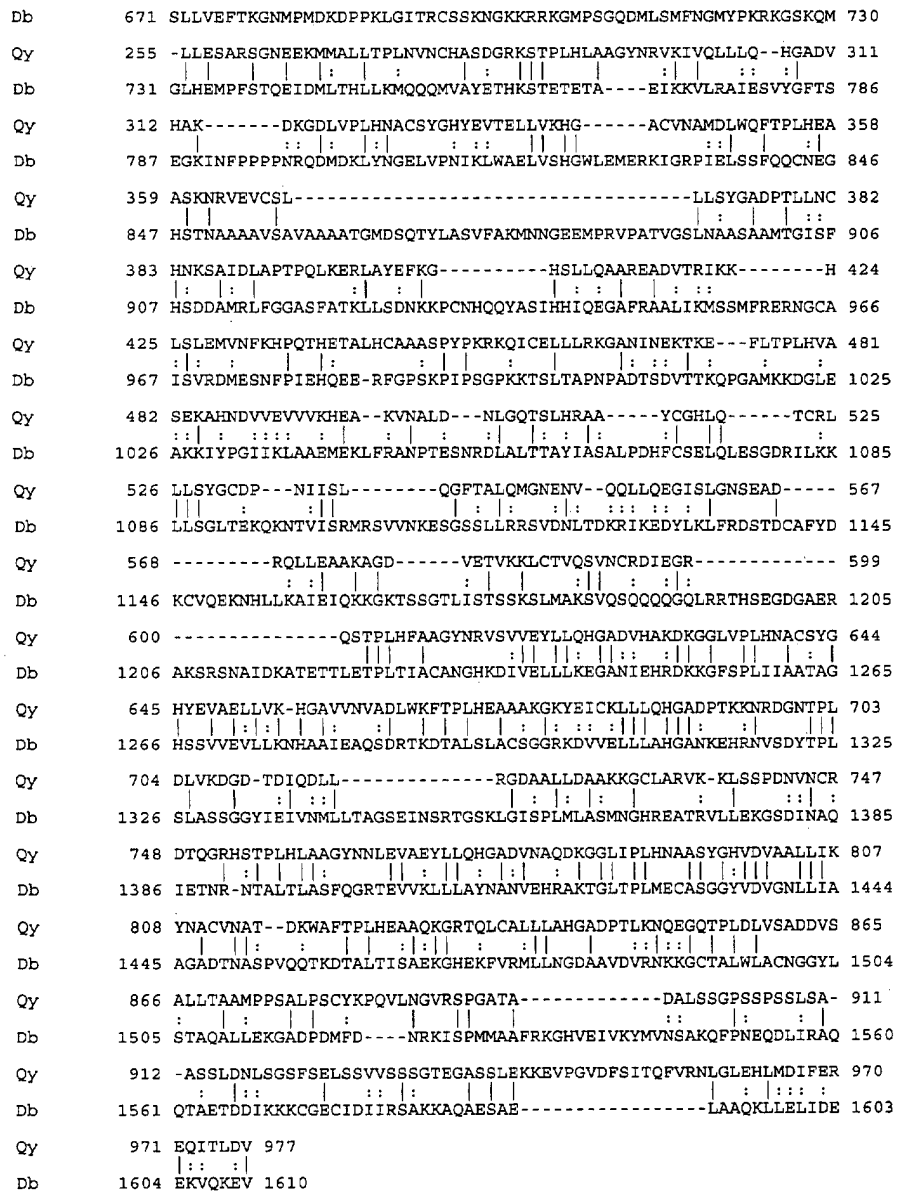
Qy          9 AGGQGAQRGARVGAAHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLA 68
              || || | |              :      |  :|    |  | ||     : :: :|:
Db        466 AGDQGDQAGV----------VKLLLSKHAEVDVSNPDTGDTPLSLAARNGYIAIMKMLI- 514

Qy         69 VAAARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDT 128
                        ||   |  :  |          || ||| :| :: ::    |   :|
Db        515 ------------EKGGDLTAGKTSPIV--------EAARNGHLECIQFILAHCKTIPQDQ 554

Qy        129 AGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHG 188
                |    |  || ||   :|| ::: ||::    |      |  |    | ||| |||  |
Db        555 LSR---ALVSAADFGSLLIVEEVIRAGADLNFEQD-ERTALMKAAKGDHFEVVQLLLSKG 610

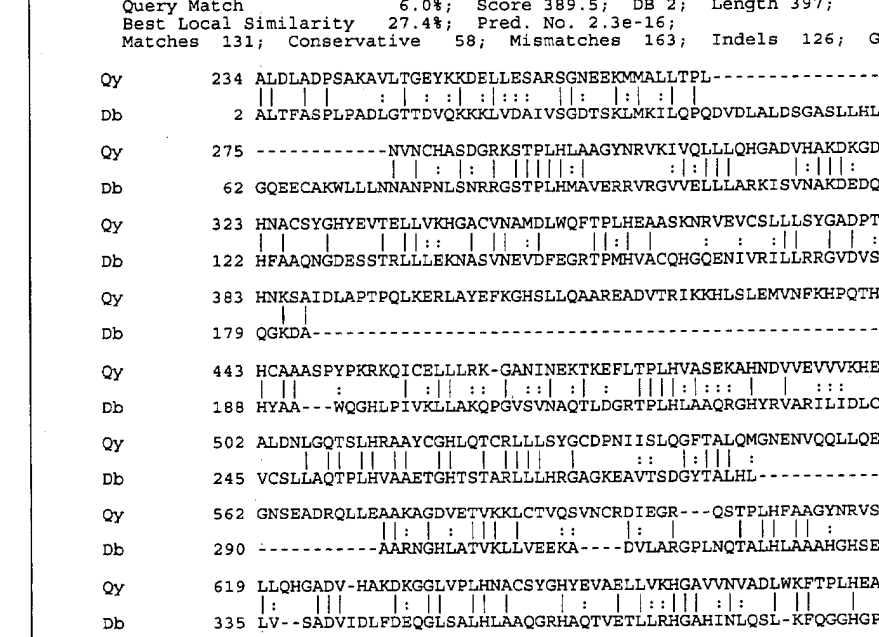
Qy        189 ADPNARDNWN-YTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG---------RTALDLA 238
              |  | : :  |  | |   :| :::    |:::||:| ::  ||           ::||
Db        611 ASVNFKSSKNDATALSLACSEGNMEIAEFLIRNGADPMLKMDDGVNCFMEVARHGSIDLM 670

Qy        239 D-----------------------------------PSAKAVLT---GEYKKDE----- 254
                                                  || : :|:   | | | :
Db        671 SLLVEFTKGNMPMDKDPPKLGITRCSSKNGKKRRKGMPSGQDMLSMFNGMYPKRKGSKQM 730
```



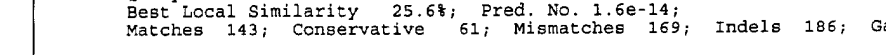
RESULT 27
T46445
hypothetical protein DKFZp434B2328.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46445
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23032
A;Accession: T46445
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-397 <AAA>
A;Cross-references: EMBL:AL137448
A;Experimental source: adult testis; clone DKFZp434B2328
C;Genetics:
A;Note: DKFZp434B2328.1

Query Match 6.0%; Score 389.5; DB 2; Length 397;
Best Local Similarity 27.4%; Pred. No. 2.3e-16;
Matches 131; Conservative 58; Mismatches 163; Indels 126; Gaps 11;



RESULT 28
T43458
hypothetical protein DKFZp434F0621.1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43458
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z22517
A;Accession: T43458
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1031 <AAA>
A;Cross-references: EMBL:AL133620
A;Experimental source: adult testis; clone DKFZp434F0621
C;Genetics:
A;Note: DKFZp434F0621.1

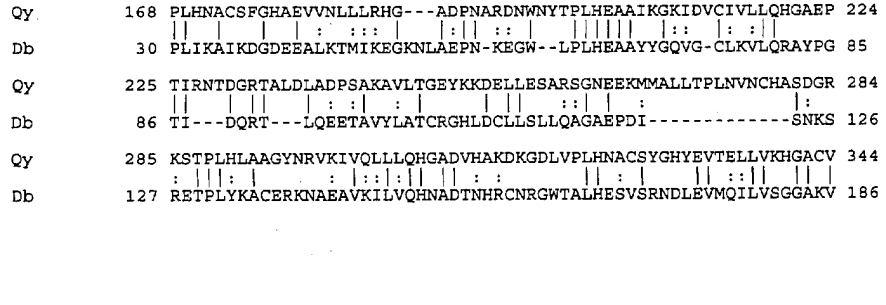
Query Match 5.7%; Score 370; DB 2; Length 1031;
Best Local Similarity 25.6%; Pred. No. 1.6e-14;
Matches 143; Conservative 61; Mismatches 169; Indels 186; Gaps 14;





RESULT 29
T46507
hypothetical protein DKFZp586M2121.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46507
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46507
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-557 <AAA>
A;Cross-references: EMBL:AL137735
A;Experimental source: adult uterus; clone DKFZp586M2121
C;Genetics:
A;Note: DKFZp586M2121.1

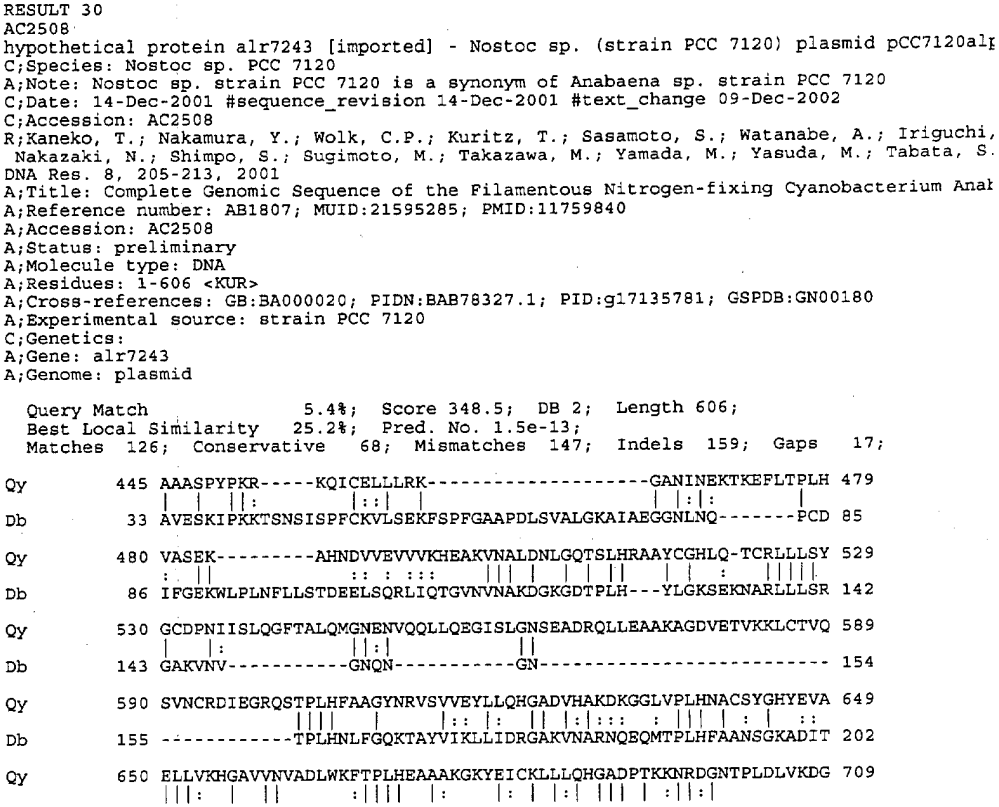
Query Match 5.4%; Score 350; DB 2; Length 557;
Best Local Similarity 24.3%; Pred. No. 1.1e-13;
Matches 133; Conservative 72; Mismatches 190; Indels 152; Gaps 15;





RESULT 30
AC2508
hypothetical protein alr7243 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alp
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AC2508
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2508
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-606 <KUR>
A;Cross-references: GB:BA000020; PIDN:BAB78327.1; PID:g17135781; GSPDB:GN00180
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7243
A;Genome: plasmid

Query Match 5.4%; Score 348.5; DB 2; Length 606;
Best Local Similarity 25.2%; Pred. No. 1.5e-13;
Matches 126; Conservative 68; Mismatches 147; Indels 159; Gaps 17;



```
Db        203 ELLISRGGDVNARSTQNWTPLHYGAS--NLEVAKKLIQAGADLTIQNREG---------- 250

Qy        710 DTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVA 769

Db        251 ----------AVIHSSSLEPAVLKL-LLDRGVNVNLRNRQGQ--TPLHIHRFQPAL--V 294

Qy        770 EYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQK 829

Db        295 KLLLARKAQVNLRDAQGKTPLYDV----NLEVARQLVAANADLNVQDNLGRTPLHQAVLE 350

Qy        830 GR----TQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQV 885

Db        351 EQSFFGSELVTLFLSKNARVDLKDKQGKTALDLAR-----------------QLNKTQI 392

Qy        886 LNGVRSPGATADALSSGPSS 905

Db        393 INLLERHMATKNSPGSGTST 412


RESULT 31
I37275
death-associated protein kinase (EC 2.7.1.-) - human
N;Alternate names: calmodulin-dependent protein kinase homolog; DAP kinase
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 10-Jul-1998
C;Accession: I37275; S39269
R;Deiss, L.P.; Feinstein, E.; Berissi, H.; Cohen, O.; Kimchi, A.
Genes Dev. 9, 15-30, 1995
A;Title: Identification of a novel serine/threonine kinase and a novel 15-kD protein as
A;Reference number: A55614; MUID:95129831; PMID:7828849
A;Accession: I37275
A;Molecule type: mRNA
A;Residues: 1-1423 <RES>
A;Cross-references: EMBL:X76104; NID:g434846; PID:g434847
C;Genetics:
A;Gene: GDB:DAPK1; DAPK
A;Cross-references: GDB:555932; OMIM:600831
A;Map position: 9q34.1-9q34.1
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A;Pathway: apoptosis
A;Note: activity is calmodulin dependent
C;Superfamily: death-associated protein kinase; ankyrin repeat homology; protein kinase
C;Keywords: apoptosis; ATP; calmodulin binding; phosphotransferase; serine/threonine-spe
F;11-267/Domain: protein kinase homology <KIN>
F;19-27/Region: protein kinase ATP-binding motif
F;285-308/Region: calmodulin binding #status predicted
F;370-402/Domain: ankyrin repeat homology <AN1>
F;403-435/Domain: ankyrin repeat homology <AN2>
F;436-468/Domain: ankyrin repeat homology <AN3>
F;470-502/Domain: ankyrin repeat homology <AN4>
F;503-535/Domain: ankyrin repeat homology <AN5>
F;536-568/Domain: ankyrin repeat homology <AN6>
F;569-601/Domain: ankyrin repeat homology <AN7>
F;602-634/Domain: ankyrin repeat homology <AN8>
F;42,64,139,141/Active site: Lys, Glu, Asp, Lys #status predicted

  Query Match             5.3%;  Score 345;  DB 1;  Length 1423;
  Best Local Similarity   22.0%;  Pred. No. 9.4e-13;
  Matches  207;  Conservative  119;  Mismatches  355;  Indels  258;  Gaps   33;

Qy        198 NYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLE 257

Db        182 NYEPLGLEADMWSIGVITYILLSGASPFLGDTKQET---LANVSA---VNYEF-EDEYFS 234

Qy        258 SARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 317

Db        235 NTSALAKDFIRRLL--------VKDPKKR-----------MTIQDSLQH-PWIKPKD-- 271

Qy        318 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 377

Db        272 --------------TQQALSRKASAVNMEKFKKF-----AARKKWKQSVRLI------- 304
```

```
Qy        378 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 437

Db        305 -----------------SLCQRLSRSFLSRSNMSVARSDDTLDEEDSFVMKAI------- 340

Qy        438 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 497

Db        341 ----IHAINDDNVPGLQHLLGSL--SNYDVNQPNKHGTPPLLIAAGCGNIQILQLLIKRG 394

Qy        498 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQE 557

Db        395 SRIDVQDKGGSNAVYWAARHGHVDTLKFLSENKCPLDVKDKSGEMALHV----------- 443

Qy        558 GISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVE 617

Db        444 ---------------AARYGHADVAQVTCAASAQIPISRTKEEETPLHCAAWHGYYSVAK 488

Qy        618 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 677

Db        489 ALCEAGCNVNIKNREGETPLLTASARGYHDIVECLAEHGADLNACDKDGHIALHLAVRRC 548

Qy        678 KYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVK 736

Db        549 QMEVIKTLLSQGCFVDYQDRHGNTPLHVACKDGNMPIV------VALCEA---NC----- 594

Qy        737 KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 796

Db        595 ------NLDISNKYGR--TPLHLAANNGILDVVRYLCLMGASVEALTTDGKTAEDLARSE 646

Qy        797 GHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 856

Db        647 QHEHVAGLLARLR---KDTHRGLFIQQLRPTQNLQPRIKLKLFGHSG-------SGKTTL 696

Qy        857 ------DLVSA------DDVSALLTAAMPPSALPSCYKPQV---LN---------GVRS- 891

Db        697 VESLKCGLLRSFFRRRRPRLSSTNSSRFPPSPLAS--KPTVSVSINNLYPGCENVSVRSR 754

Qy        892 -----PGATADALS--SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKE 944

Db        755 SMMFEPGLTKGMLEVFVAPTHHPHCSA----DDQSTK--------------AIDIQNAY 795

Qy        945 VPGV-DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN----------- 992

Db        796 LNGVGDFSVWEFSGNPVYFCCYDYFAANDPTSIHVVVFSLEEPYEIQLNPVIFWLSFLKS 855

Qy        993 --------AYGHR------------HKLIKGVERLISGQQGL----------------- 1014

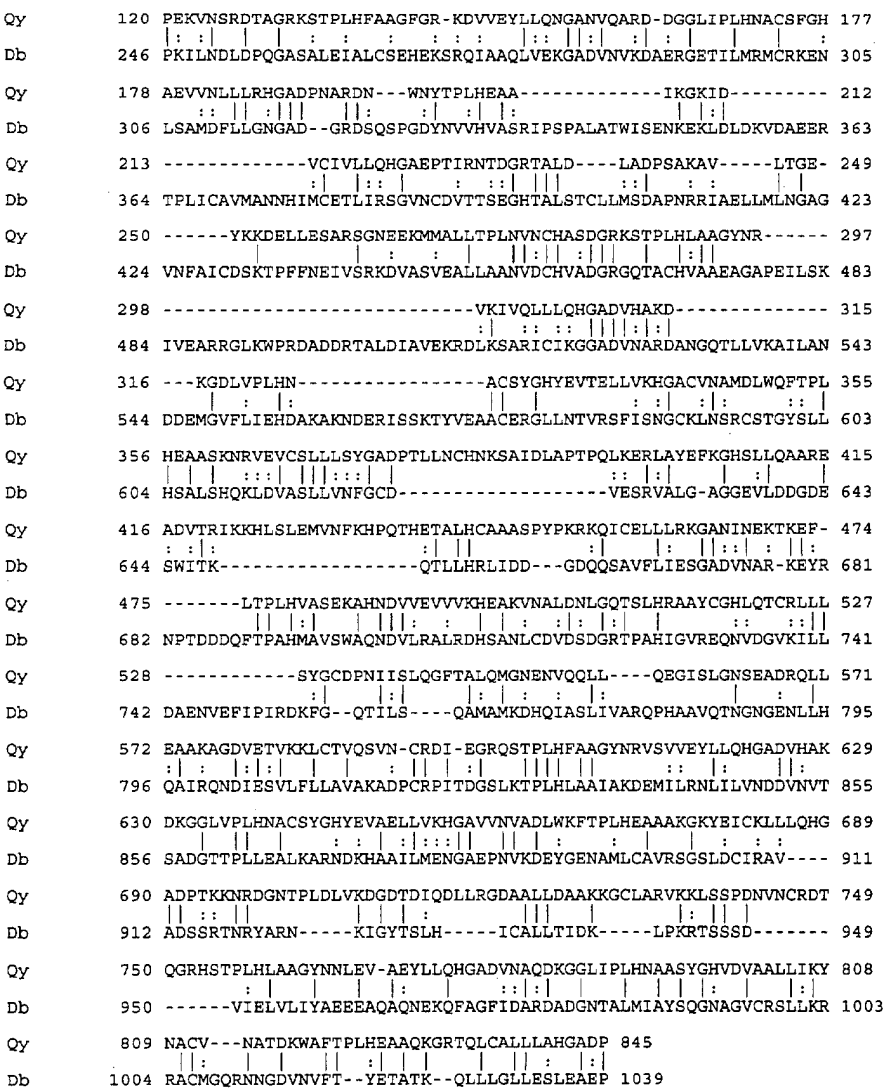
Db        856 LVPVEEPIAFGGKLKNPLQVVLVATHADIMNVPRPAGGEFGYDKDTSLLKEIRNRFGNDL 915

Qy       1015 ---NPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV 1050

Db        916 HISNKLFVLDAGASG------SKDMKVLRNHLQEIRSQI 948


RESULT 32
T21280
hypothetical protein F22G12.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21280
R;Lennard, N.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19400
A;Accession: T21280
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1107 <WIL>
A;Cross-references: EMBL:Z92831; PIDN:CAB07368.1; GSPDB:GN00019; CESP:F22G12.4
A;Experimental source: clone F22G12
C;Genetics:
A;Gene: CESP:F22G12.4
A;Map position: 1
A;Introns: 23/3; 153/2; 216/2; 268/1; 304/1; 363/2; 521/1; 562/2; 643/3; 738/3; 770/3; 8
```

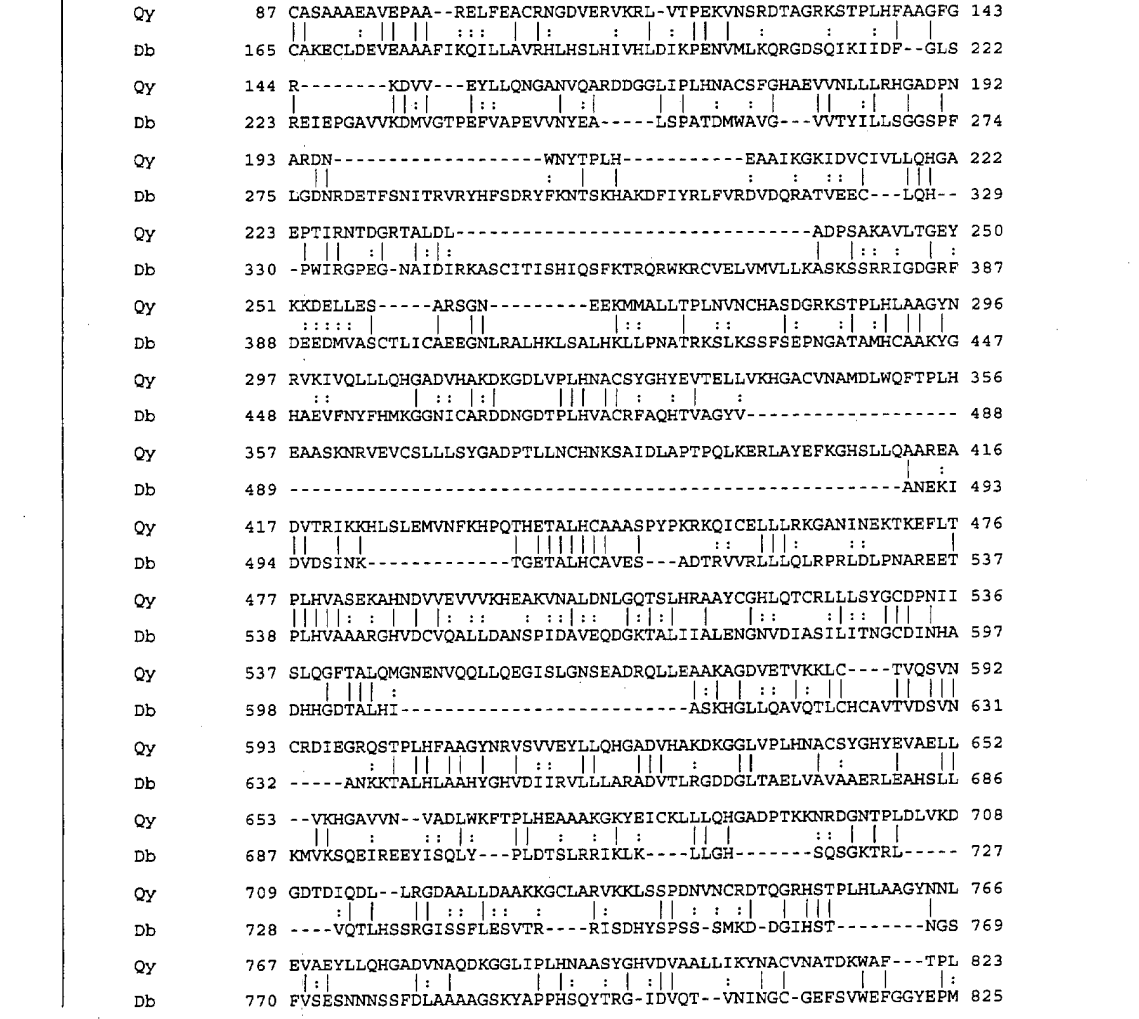
Query Match 5.3%; Score 343; DB 2; Length 1107;
Best Local Similarity 21.5%; Pred. No. 8.4e-13;
Matches 189; Conservative 126; Mismatches 325; Indels 240; Gaps 33;

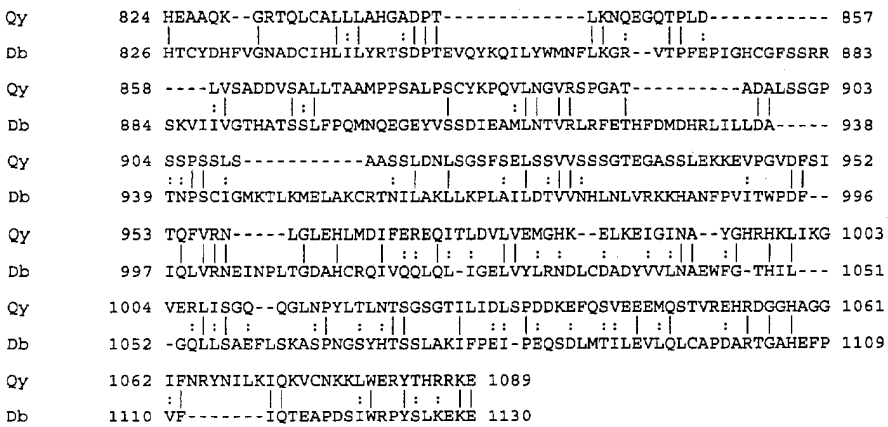


RESULT 33
T32930
hypothetical protein K12C11.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C;Accession: T32930
R;Wamsley, P.; Kramer, J.
submitted to the EMBL Data Library, January 1998

A;Description: The sequence of C. elegans cosmid K12C11.
A;Reference number: Z21248
A;Accession: T32930
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1435 <WAM>
A;Cross-references: EMBL:AF043701; PIDN:AAB97579.1; GSPDB:GN00019; CESP:K12C11.4
A;Experimental source: strain Bristol N2; clone K12C11
C;Genetics:
A;Gene: CESP:K12C11.4
A;Map position: 1
A;Introns: 79/2; 154/2; 267/2; 433/1; 533/3; 599/3; 691/3; 875/1; 1025/3; 1169/3; 1348/3
C;Superfamily: death-associated protein kinase; ankyrin repeat homology; protein kinase h

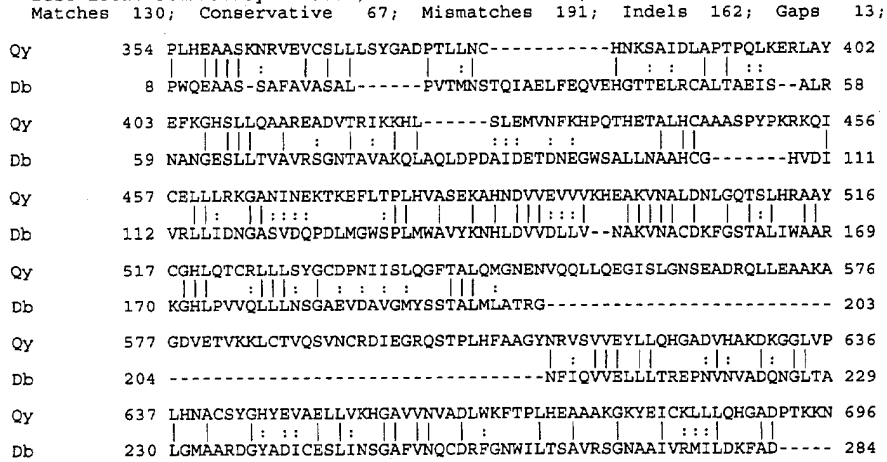
Query Match 5.1%; Score 329.5; DB 2; Length 1435;
Best Local Similarity 20.7%; Pred. No. 8.6e-12;
Matches 242; Conservative 151; Mismatches 408; Indels 367; Gaps 53;

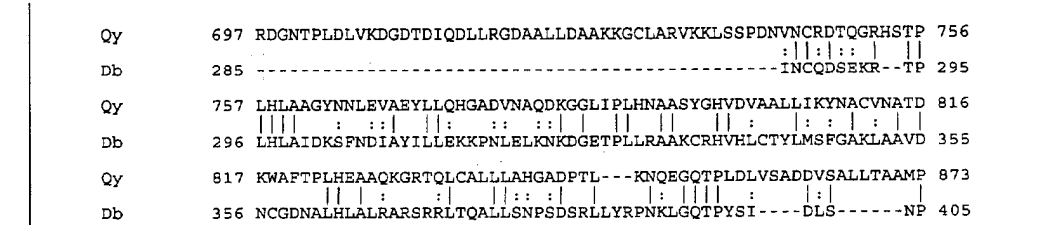




RESULT 34
T21884
hypothetical protein F36H1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21884
R;Kershaw, J.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19482
A;Accession: T21884
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1398 <WIL>
A;Cross-references: EMBL:Z68760; PIDN:CAA92996.1; GSPDB:GN00022; CESP:F36H1.2
A;Experimental source: clone F36H1
C;Genetics:
A;Gene: CESP:F36H1.2
A;Map position: 4
A;Introns: 38/3; 60/3; 94/3; 152/3; 191/3; 257/2; 290/3; 323/3; 379/1; 439/1; 485/2; 552

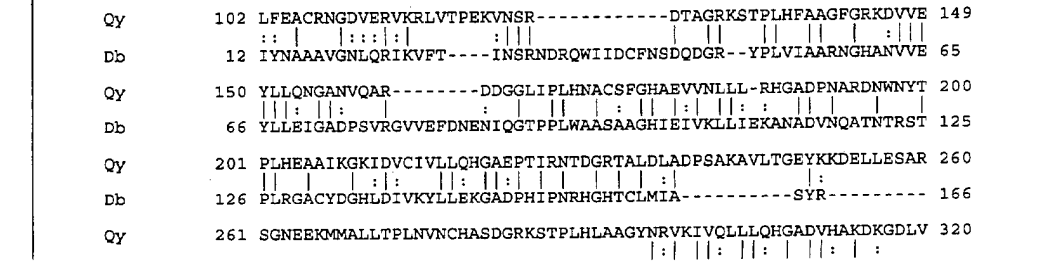
Query Match 5.0%; Score 321; DB 2; Length 1398;
Best Local Similarity 23.6%; Pred. No. 2.8e-11;
Matches 130; Conservative 67; Mismatches 191; Indels 162; Gaps 13;

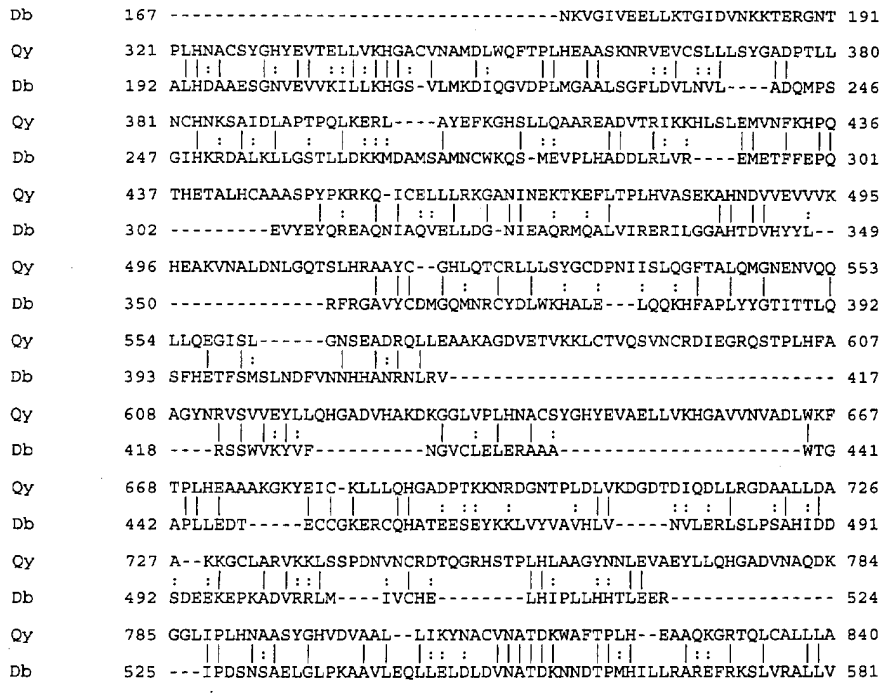




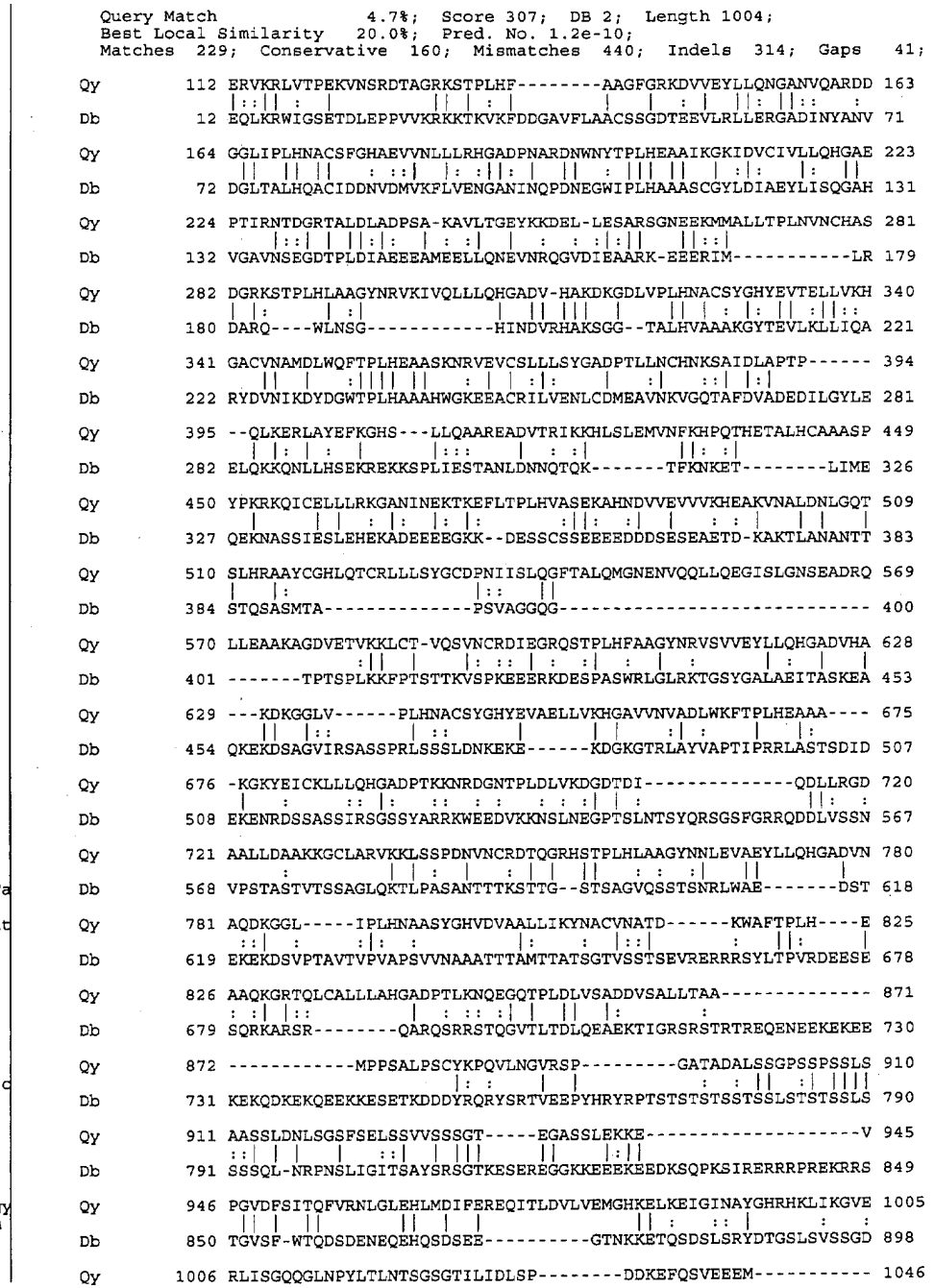
RESULT 35
A34793
sex-determining protein fem-1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 29-Oct-1999
C;Accession: A34793; T30117
R;Spence, A.M.; Coulson, A.; Hodgkin, J.
Cell 60, 981-990, 1990
A;Title: The product of fem-1, a nematode sex-determining gene, contains a motif found i
A;Reference number: A34793; MUID:90199879; PMID:2317869
A;Accession: A34793
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-656 <SPE>
A;Cross-references: GB:J03172; NID:g156309; PIDN:AAA28055.1; PID:g156310
R;Du, Z.; Le, T.T.
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of C. elegans cosmid F35D6.
A;Reference number: Z20739
A;Accession: T30117
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-656 <DUZ>
A;Cross-references: EMBL:U52001; PIDN:AAA96093.1; GSPDB:GN00022; CESP:fem-1
A;Experimental source: strain Bristol N2; clone F35D6
C;Genetics:
A;Gene: CESP:fem-1
A;Map position: 4
A;Introns: 25/3; 67/3; 88/1; 103/3; 139/1; 168/3; 290/3; 329/3; 378/2; 617/1
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;47-79/Domain: ankyrin repeat homology <AN1>
F;88-119/Domain: ankyrin repeat homology <AN2>
F;122-154/Domain: ankyrin repeat homology <AN3>
F;155-187/Domain: ankyrin repeat homology <AN4>
F;188-219/Domain: ankyrin repeat homology <AN5>
F;220-252/Domain: ankyrin repeat homology <AN6>

Query Match 4.9%; Score 315.5; DB 2; Length 656;
Best Local Similarity 23.0%; Pred. No. 1.8e-11;
Matches 184; Conservative 95; Mismatches 269; Indels 251; Gaps 31;





RESULT 36
A55142
myosin-light-chain-phosphatase (EC 3.1.3.53) large chain, long form - chicken
N;Alternate names: smooth muscle protein phosphatase 1M chain M110
C;Species: Gallus gallus (chicken)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: A55142; S51023
R;Shimizu, H.; Ito, M.; Miyahara, M.; Ichikawa, K.; Okubo, S.; Konishi, T.; Naka, M.; Ta
J. Biol. Chem. 269, 30407-30411, 1994
A;Title: Characterization of the myosin-binding subunit of smooth muscle myosin phosphat
A;Reference number: A55142; MUID:95074044; PMID:7982954
A;Accession: A55142
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1004 <SHI>
A;Cross-references: GB:D37985; NID:g633037; PIDN:BAA07201.1; PID:d1007779; PID:g633038;
R;Chen, Y.H.; Chen, M.X.; Alessi, D.R.; Campbell, D.G.; Shanahan, C.; Cohen, P.; Cohen,
FEBS Lett. 356, 51-55, 1994
A;Title: Molecular cloning of cDNA encoding the 110 kDa and 21 kDa regulatory subunits c
A;Reference number: S51022; MUID:95080427; PMID:7988720
A;Accession: S51023
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 681-1004 <CHE>
A;Cross-references: GB:S74623; NID:g802102; PIDN:AAB32730.1; PID:g802103
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: alternative splicing; phosphoric monoester hydrolase; smooth muscle; tandem
F;72-104/Domain: ankyrin repeat homology <AN1>
F;105-137/Domain: ankyrin repeat homology <AN2>



```
Db        899 RYDSAQ---------GRSGSQSYLEDRKPYCSRLEKEDSTDFKKLYEQILAENEKLKAQL 949

Qy       1047 ---------------QSTVREHRDGGHAGGIFNRYNILKIQKVCNKK----LWERYTHRR 1087
                            ::| |: |         |   ::|:::|  : |   |  : : |:
Db        950 HDTNMELTDLKLQLEKTTQRQER--------FADRSLLEMEKRVSGKSQYLLGGKKSSRK 1001

Qy       1088 KEV 1090
              |::
Db       1002 KDI 1004
```

RESULT 37
T19006
ankyrin related protein C06C3.1 - Caenorhabditis elegans
N;Contains: myosin-light-chain-phosphatase (EC 3.1.3.53)
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
C;Accession: T19006; T22086
R;Berks, M.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19058
A;Accession: T19006
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1016 <WIL>
A;Cross-references: EMBL:Z36719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:C06C3.1
A;Experimental source: clone C06C3
R;Matthews, P.
submitted to the EMBL Data Library, January 1995
A;Reference number: Z19510
A;Accession: T22086
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1016 <WI2>
A;Cross-references: EMBL:Z47809; PIDN:CAA87782.1; GSPDB:GN00020; CESP:C06C3.1
A;Experimental source: clone F42A8
C;Genetics:
A;Gene: CESP:C06C3.1
A;Map position: 2
A;Introns: 27/3; 94/3; 279/3; 352/2; 462/2; 523/3; 569/2; 657/2; 718/3; 766/3; 833/3; 89
C;Keywords: phosphoric monoester hydrolase

  Query Match 4.7%; Score 306; DB 2; Length 1016;
  Best Local Similarity 27.4%; Pred. No. 1.4e-10;
  Matches 105; Conservative 52; Mismatches 136; Indels 90; Gaps 14;

```
Qy        113 RVKRLVTPEKVNSR--------------DTAGRKSTP-LHF--------AAGFGRKDVVE 149
              : |: :|   || |              : | |   | : |        |   | :: ||
Db         13 QAKQQITFNVVNKRKDQLRRWQSSEMNAEAARRTKRPKVQFQDSDIFLSACMSGDEEEVE 72

Qy        150 YLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKG 209
               ||  |||:       ||  || :       |:|   |    ||| ||:||  :|||| || |
Db         73 ELLNKGANINTCTVDGLTALHQSVIDSKPEMVRFLCEKGADVNAQDNEGWTPLHAAAACG 132

Qy        210 KIDVCIVLLQHGAEPTIRNTDGRTALDLA-DPSAKAVLTGEYKKDEL-LESARSGNEEKM 267
              : :   | ||||: :| |:|   ||||| |   :  |  :||:  : |:: |   ::::
Db        133 NVAIVRYLCQHGADLSIVNSDKELALDLAVDEQCRDYLDDDYKRQMIDLDACR---DQEL 189

Qy        268 MALLTPLNVNCHASDGR-KSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNAC 326
              :|   :||   | |  :  | |  |                :| |
Db        190 QTMLK--DVNMWMSQGEYRDIPHHRTGG-----------------------TAMHVAA 222

Qy        327 SYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGAD--------PT 378
              |: :: |||:| |  | | |:  :|||| ||     : | :||  ||:
Db        223 GRGYTQLLELLIKAGGNVRAQDVEGWTPLHAAAHWAERDACKILLENGAELSDLTFTGAD 282

Qy        379 LLNCHNKSAIDL---------------------APTPQLKERLAYEFKGHS--LLQAAR 414
              :|   :|  ||                    ||  |:|:  :    |   :| : |
Db        283 VLGVADKECIDYLVELADTVKVQNKRKSPGSGSQPPTSILQEK-NHRMPSHEEHVLTSER 341

Qy        415 EADVTRIKKHLSLEMVNFKHPQT 437
              : |:    :|   | ::  || |
```

```
Db        342 KRDLQHKDQHSENEFLH-SHPST 363
```

RESULT 38
A59288
myosin heavy chain Myr 8 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
C;Accession: A59288
R;Patel, K.G.; Liu, C.; Cameron, P.L.; Cameron, R.S.
submitted to GenBank, November 1999
A;Description: Identification of a Novel Mammalian Myosin Class, XVI, in Developing Brai
A;Reference number: A59288
A;Accession: A59288
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1322 <PAT>
A;Cross-references: GB:AF209114; PIDN:AAF20150.1
A;Experimental source: strain Sprague-Dawley; clone KP4; cell type type 1 astrocyte
C;Superfamily: myosin motor domain homology
F;404-1132/Domain: myosin motor domain homology <MMO>

  Query Match 4.7%; Score 306; DB 2; Length 1322;
  Best Local Similarity 23.9%; Pred. No. 2.1e-10;
  Matches 151; Conservative 98; Mismatches 226; Indels 158; Gaps 24;

```
Qy        551 VQQLLQEGISLGNSEADRQLLEAAKAGDVETV---KKLCTVQSVNCRDIEGRQSTPLHF- 606
              : | | | : ||  :   :|:   :  ::     :|:    |    : |: :|  : |
Db          3 IDQCLLESLPLGQRQ---RLVRRMRCEQIKAYYEREKVFQKQEGLLKRIKPGKSQKVRFG 59

Qy        607 -------AAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 659
                    |  ::    |  ||: ||| |     |   ||    | :  :||:|:  |  |
Db         60 LADMIQDAIIHHHDKEVLQLLKEGADPHTLVSSGGSLLHLCARYDNVFIAEVLIDRGVNV 119

Qy        660 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 719
              |  |   : |:| | |    :|  ||:  ||:   :: :|| |||    :|     ||
Db        120 NHQDEDFWAPMHIACACDNPDIVLLLILAGANVLLQDVNGNIPLDYAVEGTESSAILL-- 177

Qy        720 DAALLD--AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGA 777
                | ||       | ::|                  || :      |    : |  |
Db        178 --AYLDENGVDLNSLRQIKL---------------QRPLSM------LTDVRHFLSSGG 213

Qy        778 DVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCAL 837
              ||| ::  |:  || | : |: :|  ||:::   :|  |   :|||| ||: |:| |  |
Db        214 DVNEKNDDGVTLLHMACASGYKEVVLLLLEHGGDLNGMDDGYWTPLHLAAKYGQTTLVKL 273

Qy        838 LLAHGADPTLKNQEGQTPLDLVSADDVSALLTAA------------MPPSALPSCYK-- 882
              |||| |:| | |  |: | |: ::: :  :|  |             :|  |    |:
Db        274 LLAHQANPHLVNCNGEKPSDIAASESIEEMLLKAEIAWEERMKESPSVPSLAQEELYEEI 333

Qy        883 ----PQV---LNGVRSPGATADAL------------------SSGPSSPSSLSAASSLD 916
                  |::   |: :  | |  |:|                  ||   | :|::|  :
Db        334 LHDLPELSSKLSPLVLPIAKQDSLLEKDIMFKDTTKGLCNQESQDGPPETSMVSSSSKPE 393

Qy        917 NL-------SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFE 969
              :        |   : || :  ||          | ::  | |  |  |: :: |  |:: |:
Db        394 QVQLTPPAPSDDLATLSELNDSS-----LLYEIQKRFGND-QIHTFIGDIFL--LVNPFK 445

Qy        970 REQITLDVLVEMGHKELKEIGINAYGHR---------------HKLIKGVER-----LI 1008
                |   |: :|          ::  | |               |:| :  ||       ::
Db        446 ELPIYSTVVSQM--------YLSPTGQRSPSLPPHLFSCAERAFHRLFQ--ERRPQNIIL 495

Qy       1009 SGQQG---------LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHA 1059
              ||::|          :  :||   | | |:        |  |:     :::         |||
Db        496 SGERGSGKTQASKQIMKHLTSRASSSCTMF------DSRFKHAICILEAF-------GHA 542

Qy       1060 GGIFNRYNILKIQKVCNKKLWE-RYTHRRKEVS 1091
                |  :   ||        || ::  ||| |:
Db        543 KTTLNNVSSCLIQ------YWELQFCQRRKHVT 569
```

RESULT 39
B45771

```
2-5A-dependent RNAase - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
C;Accession: B45771
R;Zhou, A.; Hassel, B.A.; Silverman, R.H.
Cell 72, 753-765, 1993
A;Title: Expression cloning of 2-5A-dependent RNAase: a uniquely regulated mediator of i
A;Reference number: A45771; MUID:93201598; PMID:7680958
A;Accession: B45771
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-679 <ZHO>
A;Cross-references: GB:L10382; NID:g293264; PIDN:AAA37117.1; PID:g293265
A;Note: sequence extracted from NCBI backbone (NCBIP:127769)
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;58-90/Domain: ankyrin repeat homology <AN22>
F;124-156/Domain: ankyrin repeat homology <AN21>

  Query Match             4.7%;  Score 305;  DB 2;  Length 679;
  Best Local Similarity   27.5%;  Pred. No. 8.6e-11;
  Matches  148;  Conservative   69;  Mismatches  232;  Indels   90;  Gaps   23;

Qy        565 EADRQLLEAAKAGDVETVKKLCT--VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 622
Db         24 EDDSSLIKAVQKGDVVRVQQLLEKGADANACEDTWG--WTPLHNAVQAGRVDIVNLLLSH  81

Qy        623 GADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEIC 682
Db         82 GADPHRRKKNGATPFIIAGIQGDVKLLEILLSCGADVNECDENGFTAFMEAAERGNAEAL 141

Qy        683 KLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDA-ALLDAAKKGCL--ARVKKLS 739
Db        142 RFLFAKGAN-----------VNLRRQTTKDKRRLKQGGATALMSAAEKGHLEVLRILLND 190

Qy        740 SPDNVNCRDTQGRHS---TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 796
Db        191 MKAEVDARDNMGRNALIRTLLNWDC-ENVEEITSILIQHGADVNVRGERGKTPLIAAVER 249

Qy        797 GHVDVAALLIKYNAC-VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 855
Db        250 KHTGLVQMLLSREGINIDARDNEGKTALLIAVDKQLKEIVQLLLEKGAD----------- 298

Qy        856 LDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSL 915
Db        299 ----KCDDL-VWIARRNHDYHLVKLLLPYVANPDTDPPA-------GDWSPHSSRWGTAL 346

Qy        916 DNLSG----SFSELSSVVSS----SGT-EGASSL---EKKEVPGVDF--------SITQF 955
Db        347 KSLHSMTRPMIGKLKIFIHDDYKIAGTSEGAVYLGIYDNREVAVKVFRENSPRGCKEVSC 406

Qy        956 VRNLGLEH--LMDIFERE------------QITLDVLVEMGHKELKEIGINAYGHR--H 998
Db        407 LRDCG-DHSNLVAFYGREDDKGCLYVCVSLCEWTLEEFLRLPREEPVENGEDKFAHSILL 465

Qy        999 KLIKGVERL-ISG--QQGLNPYLTLNTSGSGTILIDLSPDDK---EFQSVEEEMQSTVR 1051
Db        466 SIFEGVQKLHLHGYSHQDLQPQNILIDSKKAVRLADFDQSIRWMGESQMVRRDLEDLGR 524


RESULT 40
T27499
hypothetical protein ZC15.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27499
R;Basham, V.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z20377
A;Accession: T27499
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-633 <WIL>
A;Cross-references: EMBL:Z93396; NID:e1188462; PIDN:CAB07714.1; GSPDB:GN00023; CESP:ZC15
A;Experimental source: clone ZC15
C;Genetics:
A;Gene: CESP:ZC15.7
A;Map position: 5
A;Introns: 15/3; 69/2; 100/3; 148/3; 194/1; 305/1; 404/3; 520/3

  Query Match             4.6%;  Score 299.5;  DB 2;  Length 633;
  Best Local Similarity   23.8%;  Pred. No. 1.7e-10;
  Matches  161;  Conservative   90;  Mismatches  257;  Indels  169;  Gaps   26;

Qy        202 LHEAAIK-GKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
Db         25 LEKQAIKNGSTDV-----DHFWE--ILDDKGRNVL------FHAAITDNLKNFQFIMKTR 71

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQ-------------H 307
Db         72 RVKKEALLRL----------TDYNKATVLHWATQYNCKQVVKEIVQTFDKLNPEDLKILH 121

Qy        308 GADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQF-------TPLHEAAS 360
Db        122 GL-ILAKDSESVTPLHIAATKQDTKILKIFVEILKTPKVMELFSIVKDKRDRSPLHYAAC 180

Qy        361 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYE----FKGHSLLQAAREA 416
Db        181 KVNLEALRILLF--VDP-------NGGPDFGFTVDQRDKFGIIPLMCAVGVNLPQAIPVI 231

Qy        417 DVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 476
Db        232 RFLEKKKPVSKTRQN----RDGMTALHIAVAA---RNLEAVQLLIELGSSVDLVDNEQRT 284

Qy        477 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNII 536
Db        285 PLHYAAEQGYPEIVKFLLCNGARNSTRDHIGATPAHYAA---------------------- 323

Qy        537 SLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDI 596
Db        324 --------QFSVECLKILFAES---------------------------KITEVN--DN 345

Qy        597 EGRQSTPLHFAAGYNRVSVVEYLLQH--GADVHAKDKGGLVPLHNACSYGHYEVAELLVK 654
Db        346 EGRSC--LMWAVCAGNIEVINYLIQREDAPKRAACDKNGYTALHLAAMVGHEKVCKILTN 403

Qy        655 ---------HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL 705
Db        404 QPTSIFPYFQGWSLSERDNHSNTALHLASGRGHTDVLRCLVASGANMNDVDEVGRTAV-- 461

Qy        706 VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLS----SPDNVNCR---DTQGRHSTPLH 758
Db        462 -------FWACMGGQAHTLH-----CM--IKELGFEWRTSGSVNTRPKCDKNGR--TALH 505

Qy        759 LAAGYNNLEVAEYLLQHGAD-------VNAQDKGGLIPLHNAASYGHVDVAALLIKYNAC 811
Db        506 AAAVESSACINVLLSIEKEDNFLSSPLVGWLDKNGETALHEACLASRIDCILSLLNGGSA 565

Qy        812 VNATDKWAFTPLHEAAQ 828
Db        566 VNAFGRPQRTPLDCAMQ 582


RESULT 41
D84448
probable ankyrin [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 08-Sep-2002
C;Accession: D84448
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84448
A;Status: preliminary
```

```
A;Molecule type: DNA
A;Residues: 1-247 <STO>
A;Cross-references: GB:AE002093; NID:g4335756; PIDN:AAD17433.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g03430
A;Map position: 2
C;Superfamily: hypothetical protein YGR232w; ankyrin repeat homology

  Query Match             4.6%;  Score 299;  DB 2;  Length 247;
  Best Local Similarity   30.7%;  Pred. No. 4.1e-11;
  Matches   81;  Conservative   38;  Mismatches   87;  Indels   58;  Gaps    5;

Qy        541 FTALQMGNENV-----QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQS----V 591
Db         17 FKAAEWGDSSLFMSLSEEQLSKSLNFRNEDGRSLLHVAASFGHSQIVKLLSSSDEAKTVI 76

Qy        592 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 651
Db         77 NSKDDEG--WAPLHSAASIGNAELVEVLLTRGADVNAKNNGGRTALHYAASKGRLEIAQL 134

Qy        652 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 711
Db        135 LLTHGAKINITDKVGCTPLHRAASVGKLEVCEFLIEEGAEIDATDKMGQTAL-------- 186

Qy        712 DIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEY 771
Db        187 ----------------------------MHSVICDDKQLKVSMD----------QVAFL 207

Qy        772 LLQHGADVNAQDKGGLIPLHNAAS 795
Db        208 LIRHGADVDVEDKEGYTVLGRATN 231


RESULT 42
S30385
G9a protein - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jan-2000
C;Accession: S30385
R;Milner, C.M.; Campbell, R.D.
Biochem. J. 290, 811-818, 1993
A;Title: The G9a gene in the human major histocompatibility complex encodes a novel prot
A;Reference number: S30385; MUID:93207535; PMID:8457211
A;Accession: S30385
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1001 <MIL>
A;Cross-references: EMBL:X69838; NID:g287864; PIDN:CAA49491.1; PID:g287865
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;641-673/Domain: ankyrin repeat homology <ANR>

  Query Match             4.6%;  Score 297;  DB 2;  Length 1001;
  Best Local Similarity   23.4%;  Pred. No. 4.9e-10;
  Matches  143;  Conservative   93;  Mismatches  227;  Indels  148;  Gaps   22;

Qy         83 GGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHFAAGF 142
Db        406 GG--CLSAVGLPLGPGREALEKALVIQESERRKKL----RFHPRQ--------LYLSVKQ 451

Qy        143 GR-KDVVEYLLQN-GANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200
Db        452 GELQKVILMLLDNLDPNFQSDQQSKRTPLHAAAQKGSVEICHVLLQAGANINAVDKQQRT 511

Qy        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260
Db        512 PLMEAVVNNHLEVARYMVQRGGCVYSKEEDGSTCLH-------------------HAAK 551

Qy        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320
Db        552 IGNLEMVSLLLSTGQVDVNAQDSGGWTPIIWAAEHKHIEVIRMLLTRGADVTLTDNEENI 611

Qy        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380
Db        612 CLHWASFTGSAAIAEVLLNARCDLHAVNYHGDTPLHIAARESYHDCVLLFLSRGANPELR 671

Qy        381 NCHNKSAIDLAP-------TPQLKERLAYEFKGHSLLQAARE--ADVTRIKKHLSLEMVN 431
Db        672 NKEGDTAWDLTPERSDVWFALQLNRKLRLGV-GNRAIRTEKIICRDVARGYENVPIPCVN 730

Qy        432 ----------FKH-PQTHETAL----------HCAAASPYPKRKQIC-ELLLR----KG 464
Db        731 GVDGEPCPEDYKYISENCETSTMNIDRNITHLQHCTCVDDCSSSNCLCGQLSIRRWYDKD 790

Qy        465 ANINEKTKEFLTPLHVASEKA-------HNDVV-------------------------- 490
Db        791 GRLLQEFNKIEPPLIFECNQACSCWRNCKNRVVQSGIKVRLQLYRTAKMGWGVRALQTIP 850

Qy        491 ----------EVVVKHEAKVNA-------LDNL-GQTSLHRAAYCGHLQTCRLLLSYGCD 532
Db        851 QGTFICEYVGELISDAEADVREDDSYLFDLDNKDGEVYCIDARYYGNISR---FINHLCD 907

Qy        533 PNIISLQGFTALQMGNENVQ---------QLLQEGISLGNSEADRQLLEAAKAGDVETVK 583
Db        908 PNIIPVRVF----MLHQDLRFPRIAFFSSRDIRTGEELGFDYGDR-------FWDIKSKY 956

Qy        584 KLCTVQSVNCR 594
Db        957 FTCQCGSEKCK 967


RESULT 43
T21720
hypothetical protein F34D10.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21720
R;Kershaw, J.
submitted to the EMBL Data Library, June 1994
A;Reference number: Z19464
A;Accession: T21720
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1295 <WIL>
A;Cross-references: EMBL:Z34799; PIDN:CAA84316.1; GSPDB:GN00021; CESP:F34D10.6
A;Experimental source: clone F34D10
C;Genetics:
A;Gene: CESP:F34D10.6
A;Map position: 3
A;Introns: 7/3; 46/3; 109/1; 204/2; 227/3; 266/3; 298/2; 336/1; 397/1; 429/3; 512/2; 597

  Query Match             4.6%;  Score 297;  DB 2;  Length 1295;
  Best Local Similarity   22.9%;  Pred. No. 7.4e-10;
  Matches  162;  Conservative  109;  Mismatches  253;  Indels  184;  Gaps   32;

Qy         97 PAARELFEACRNGDVERVKRLVTPEK---VNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQ 153
Db        599 PSEEQVEEWLATGEVLKLEQIVLDGRGHMLREKKTVNASSN--EFLSG-----LTQYLTK 651

Qy        154 NGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDV 213
Db        652 IDA------------IHKAVEEGDVRRVKSLIDRPLLSTARDNYGMTPIHKALLHGQTNT 699

Qy        214 CIVLLQHGAEPTIRN-TDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 272
Db        700 VRFLL--GRFPSCVNATDHRIFGILVD----SVSNGVFNKSPLKKNA------------- 740

Qy        273 PLNVNCHASDGRKSTPLHLAA----GYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSY 328
Db        741 ----------GR--TPLHYAAADPNGEHMIKVLQ---KSGGDAFIEDKHGHTPFYYRTHG 785

Qy        329 GHYEVTELLVKHGACVNAMDLWQFT-PLHEAASK---------NRVEVCSLLLSYGADPT 378
Db        786 QRLNVRSM--KDNAVMNQLISGQLSRPLLQDLEEDIYDWIHTGNIGKLEELVLTGYADLL 843

Qy        379 LLNCHNKSAID----LAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 434
```

```
Db        844 LGRNHEVDDADSIGFLEVLPQ------YQAKIQSIHKAVETGNLRAVKL---------- 886

Qy        435 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 494
                                  |  ||  ::  |:  |:||| |      |: : ::
Db        887 ----------------------LTDRKKLSLCRDTRG-LSPLHKAIVFERTDIAKYLI 921

Qy        495 K-HEAKVNALDNLGQTSLHRAAYC---GHLQTCRLLLSYGCDPNIISLQG---------- 540
              : :   |||:|   :| || ||     |:|   :::   | ||||    |
Db        922 RNYPQSVNAMDQKKRTPLHYAAALKDGGYLY--KVMRKSGADPNIYDCNGRPAKYYLKHT 979

Qy        541 ----FTALQMGNE-NVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD 595
                  :|:::     ::|:|   ::    |:  |  :  : | :  :::|       | |
Db        980 GEIDLSAMRLDTRAALKQVLHNRVAPSYLESSIQ--QWLRDGQLAKLEQLVL---SGCGD 1034

Qy        596 -IEGRQST-------------------PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLV 635
               :: | ||                    :| :     :  |: |::      |:|: |
Db       1035 LLQSRTSTHTETQAFLDRLPDYMDKIDGIHKSIKEGNLDKVKELMKTKKLAIARDRFGCT 1094

Qy        636 PLHNACSYGHYEVAELLVKH-GAVVNVADLWKFTPLHEAAAK---GKYEICKLLLQHGAD 691
              |||:|  : | |:   :  | :|:|  |  | | :| |||    | |   |:| : |||
Db       1095 PLHSAVVHEHTEIVRYIAGHYNSVLNAPDYNKRTAMHYAAAARDGGHY--LKILGKAGAD 1152

Qy        692 PTKKNRDGNTP----------LDLVKDGDTD---IQDLLRGDAALLDA 726
              |   : :| ||          | |:|| | |   | :    |  |:|:
Db       1153 PMAVDNEGRTPDYYRRNAVIDLKLIKDRDEDFETINEEYLEDGPLIDS 1200
```

RESULT 44
A42024
transcription factor NF-kappa-B2, p100 splice form - human
N;Alternate names: lyt-10; major histocompatibility complex class I-specific transcripti
in NFKB2; transcription factor H2TF1
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Aug-2002
C;Accession: A42024; A41645; A48525
R;Bours, V.; Burd, P.R.; Brown, K.; Villalobos, J.; Park, S.; Ryseck, R.P.; Bravo, R.; K
Mol. Cell. Biol. 12, 685-695, 1992
A;Title: A novel mitogen-inducible gene product related to p50/p105-NF-kappa B participa
A;Reference number: A42024; MUID:92123193; PMID:1531086
A;Accession: A42024
A;Molecule type: mRNA
A;Residues: 1-900 <BOU>
A;Cross-references: GB:S76638; NID:g243420; PIDN:AAB21124.1; PID:g243421
A;Experimental source: peripheral blood T-cells
A;Note: sequence extracted from NCBI backbone (NCBIN:76638, NCBIP:76639)
R;Neri, A.; Chang, C.C.; Lombardi, L.; Salina, M.; Corradini, P.; Maiolo, A.T.; Chaganti
Cell 67, 1075-1087, 1991
A;Title: B cell lymphoma-associated chromosomal translocation involves candidate oncogen
A;Reference number: A41645; MUID:92103674; PMID:1760839
A;Accession: A41645
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-42,'A',44-143,'K',145-432,'EP',435-454,'HG',457-469,'A',471-663,'G',665-74
R;Potter, D.A.; Larson, C.J.; Eckes, P.; Schmid, R.M.; Nabel, G.J.; Verdine, G.L.; Sharp
J. Biol. Chem. 268, 18882-18890, 1993
A;Title: Purification of the major histocompatibility complex class I transcription fact
A;Reference number: A48525; MUID:93366804; PMID:8360178
A;Accession: A48525
A;Molecule type: protein
A;Residues: 'K',460-461,'X',463-469,'A';'KX',581-597;'KX',637-638,'X',640-645,'X',647 <P
A;Experimental source: HeLa cells
A;Note: sequence modified after extraction from NCBI backbone
C;Comment: For other splice forms, see PIR:S17233, PIR:I38609, PIR:A57034.
C;Comment: The full-length form of this protein is the MHC class I-specific transcriptic
.
C;Genetics:
A;Gene: GDB:NFKB2
A;Cross-references: GDB:128804; OMIM:164012
A;Map position: 10q24-10q24
C;Superfamily: human transcription factor NF-kappa-B2; ankyrin repeat homology; rel homo
C;Keywords: alternative splicing; DNA binding; nucleus; transcription factor; transcript
F;38-342/Domain: rel homology <REL>
F;310-313/Region: nuclear location signal
F;337-341/Region: nuclear location signal

```
  Query Match             4.6%;  Score 294.5;  DB 2;  Length 900;
  Best Local Similarity   26.9%;  Pred. No. 5.9e-10;
  Matches  133;  Conservative   58;  Mismatches  182;  Indels  121;  Gaps   19;

Qy        546 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 605
              :   : : ||  |::      |||: | | :       : | | |  |      ||||
Db        453 LAQRSARALLDYGVT-----ADRRALLAGQ-------RHLLTAQDEN-------GDTPLH 493

Qy        606 FAAGYNRVSVVE---YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 662
              |  : : ||:|   |:: |  |:                                |||:
Db        494 LAIIHGQTSVIEQIVYVIHHAQDL----------------------------GVVNLT 523

Qy        663 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAA 722
              :   |||| |   |:  :   ||: ||||   :| |:: : |          :|||   |
Db        524 NHLHQTPLHLAVITGQTSVVSFLLRVGADPALLDRHGDSAMHLALRAGAGAPELLR---A 580

Qy        723 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 782
              || :           : |  ::  | :| :  |:|||    : |  : |:  ||:| |
Db        581 LLQSGAP---------AVPQLLHMPDFEGLY--PVHLAVRARSPECLDLLVDSGAEVEAT 629

Qy        783 DK-GGLIPLHNAASYGHVD-VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 840
              :: ||   || |     :  |  |: |  | |||      |||| ||  |   |  |||
Db        630 ERQGGRTALHLATEMEELGLVTHLVTKLRANVNARTFAGNTPLHLAAGLGYPTLTRLLLK 689

Qy        841 HGADPTLKNQE----------------------------GQTPLDLVSADDVSALLTAA 871
              |||   :|:|                            | |||||  :  |  ||  |
Db        690 AGADIHAENEEPLCPLPSPPTSDSDSDSEGPEKDTRSSFRGHTPLDLTCSTKVKTLLLNA 749

Qy        872 M-----PPSALPSCYKPQVLNGVRSPGATA----DALSSGPSSPSSLSAASSLDNLSGSF 922
                  ||   ||  | :     | | ||    : |  || :  | :  :    |
Db        750 AQNTMEPPLTPPSPAGPGL-----SLGDTALQNLEQLLDGPEAQGSWAELAERLGLRSLV 804

Qy        923 SELSSVVSSSGTEGASSLEKKEVPGVDFS-ITQFVRNLGLEHLMDIFEREQITLDVLVEM 981
                | ||   | |   |: | | : : : : ::|||  : :  |   | | |
Db        805 DTYRQTTSPSG----SLLRSYELAGGDLAGLLEALSDMGLEEGVRLL-RGPETRDKLPST 859

Qy        982 GHKELKEIGINAYG 995
                |:||   :|||
Db        860 --TEVKE--DSAYG 869
```

RESULT 45
S68418
protein phosphatase 1M chain M110 isoform - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Sep-1999
C;Accession: S68418
R;Haystead, C.M.M.; Gailly, P.; Somlyo, A.P.; Somlyo, A.V.; Haystead, T.A.J.
FEBS Lett. 377, 123-127, 1995
A;Title: Molecular cloning and functional expression of a recombinant 72.5 kDa fragment c
A;Reference number: S68418; MUID:96128211; PMID:8543033
A;Accession: S68418
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-658 <HAY>
A;Cross-references: GB:U50185; NID:g1236649; PIDN:AAA92961.1; PID:g1236650
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;62-94/Domain: ankyrin repeat homology <AN1>

```
  Query Match             4.4%;  Score 285.5;  DB 2;  Length 658;
  Best Local Similarity   29.3%;  Pred. No. 1.3e-09;
  Matches  100;  Conservative   54;  Mismatches  142;  Indels   45;  Gaps   10;

Qy        112 ERVKRLVTPEKVNSRDTAGRKSTPLHF--------AAGFGRKDVVEYLLQNGANVQARDD 163
              |::|| :  |       |: | : |         |   |  | |  ||  ||::   :
Db          2 EQLKRWIGSETDLEPPVVKRQKTKVKFDDGAVFLAACSSGDTDEVLKLLHRGADINYANV 61

Qy        164 GGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 223
              ||  || ||   : ::|  |: :||: |  ||  : ||| ||  | :|:   |: ||
Db         62 DGLTALHQACIDDNVDMVKFLVENGANINQPDNEGWIPLHAAASCGYLDIAEFLIGQGAH 121
```

```
Qy        224 PTIRNTDGRTALDLADPSA-KAVLTGEYKKDEL-LESARSGNEEKMMALLTPLNVNCHAS 281
              |::| | ||:|:  | : :|  |  :  : :|:||   ||::|
Db        122 VGAVNSEGDTPLDIAEEEAMEELLQNEVNRQGVDIEAARK-EEERIM----------LR 169

Qy        282 DGRKSTPLHLAAGYNRVKIVQLLLQHGADV-HAKDKGDLVPLHNACSYGHYEVTELLVKH 340
              | |:     | :|            | :|| |||  |    || | : |: || :||::
Db        170 DARQ----WLNSG-----------HISDVRHAKSGG--TALHVAAAKGYTEVLKLLIQA 211

Qy        341 GACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTQLKERL 400
              |  ||  |   :|||| ||   : | | :|:    |   :|   ::| |:|    |
Db        212 GYDVNIKDYDGWTPLHAAAHWGKEEACRILVDNLCDMETVNKVGQTAFDVADEDILGYLE 271

Qy        401 AYEFKGHSLLQAAREADVTRIKKHLSLE----MVNFKHPQT 437
              : | : |    |:     |:   ::|      |||: :|
Db        272 ELQKKQNLLHSEKRDKKSPLIESTANMENNQPQKNFKNKET 312

Search completed: March 25, 2004, 16:45:12
Job time : 43.3162 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:37:05 ; Search time 13.1624 Seconds
(without alignments)
4351.577 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5766	99.9	1166	1	TNK2_HUMAN	Q9h2k2 homo sapien
2	4889.5	84.8	1327	1	TNK1_HUMAN	O95271 homo sapien
3	875.5	15.2	4377	1	ANK3_HUMAN	Q12955 homo sapien
4	860.5	14.9	3924	1	ANK2_HUMAN	Q01484 homo sapien
5	842	14.6	1880	1	ANK1_HUMAN	P16157 homo sapien
6	837.5	14.5	1862	1	ANK1_MOUSE	Q02357 mus musculu
7	707.5	12.3	1059	1	Y379_HUMAN	O15084 homo sapien
8	571	9.9	768	1	YB23_HUMAN	Q9ulj7 homo sapien
9	464.5	8.1	1401	1	LATA_LATMA	P23631 latrodectus
10	449.5	7.8	1169	1	AFY1_MOUSE	Q810b6 mus musculu
11	439	7.6	1169	1	AFY1_HUMAN	Q9p2r3 homo sapien
12	420.5	7.3	603	1	V162_FOWPV	Q9j569 fowlpox vir
13	394.5	6.8	747	1	V222_FOWPV	Q9j513 fowlpox vir
14	390.5	6.8	832	1	ANR3_HUMAN	P57078 homo sapien
15	380.5	6.6	668	1	V244_FOWPV	Q9j4z6 fowlpox vir
16	370	6.4	1432	1	DAK1_HUMAN	P53355 homo sapien
17	368	6.4	596	1	V024_FOWPV	Q9j5h7 fowlpox vir
18	364	6.3	1442	1	DAK1_MOUSE	Q80ye7 mus musculu
19	351	6.1	542	1	V155_FOWPV	Q9j5a7 fowlpox vir
20	350	6.1	587	1	ASB2_HUMAN	Q96q27 homo sapien
21	342	5.9	1150	1	ANS1_MOUSE	P59672 mus musculu
22	334.5	5.8	775	1	ANR5_MOUSE	Q9d2j7 mus musculu
23	330	5.7	1267	1	HMT1_HUMAN	Q9h9b1 homo sapien
24	327.5	5.7	525	1	V228_FOWPV	Q9j507 fowlpox vir
25	322.5	5.6	776	1	ANR5_HUMAN	Q9nu02 homo sapien
26	322	5.6	525	1	ASB3_MOUSE	Q9wv72 mus musculu
27	321	5.6	436	1	V245_FOWPV	Q9j4z5 fowlpox vir
28	319	5.5	592	1	V246_FOWPV	Q9j4z4 fowlpox vir
29	317	5.5	1134	1	ANS1_HUMAN	Q92625 homo sapien
30	310	5.4	528	1	PPGA_HUMAN	Q96i34 homo sapien
31	309	5.4	524	1	PPGA_MOUSE	Q923m0 mus musculu
32	309	5.4	568	1	PPGB_MOUSE	Q8vhq3 mus musculu
33	306.5	5.3	518	1	ASB3_HUMAN	Q9y575 homo sapien
34	305	5.3	735	1	RN5A_MOUSE	Q05921 mus musculu
35	303.5	5.3	568	1	PPGB_BOVIN	Q95n27 bos taurus
36	301.5	5.2	567	1	PPGB_HUMAN	Q96t49 homo sapien
37	299.5	5.2	692	1	ANR6_HUMAN	Q9y2g4 homo sapien
38	295.5	5.1	360	1	Y045_METMA	Q8q0u0 methanosarc
39	291.5	5.1	950	1	ORP1_HUMAN	Q9bxw6 homo sapien
40	289	5.0	899	1	KBF2_HUMAN	Q00653 homo sapien
41	287.5	5.0	656	1	FEM1_CAEEL	P17221 caenorhabdi
42	286.5	5.0	741	1	RN5A_HUMAN	Q05823 homo sapien
43	285	4.9	461	1	V218_FOWPV	Q9j517 fowlpox vir
44	285	4.9	1083	1	HOS4_YEAST	P40480 saccharomyc
45	282.5	4.9	1210	1	BAT8_HUMAN	Q96kq7 homo sapien

ALIGNMENTS

RESULT 1
TNK2_HUMAN
ID TNK2_HUMAN STANDARD; PRT; 1166 AA.
AC Q9H2K2; Q9H8F2; Q9HAS4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tankyrase 2 (EC 2.4.2.30) (TANK2) (Tankyrase II) (TNKS-2) (TRF1-
DE interacting ankyrin-related ADP-ribose polymerase 2) (Tankyrase-like
DE protein) (Tankyrase-related protein).
GN TNKS2 OR TNKL OR TANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=21072550; PubMed=11205898;
RA Monz D., Munnia A., Comtesse N., Fischer U., Steudel W.-I., Feiden W.,
RA Glass B., Meese E.U.;
RT "Novel tankyrase-related gene detected with meningioma-specific
RT sera.";
RL Clin. Cancer Res. 7:113-119(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast carcinoma;
RX MEDLINE=21190090; PubMed=11294570;
RA Kuimov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,
RA Scanlan M.J., Jongeneel C.V., Lagarkova M.A., Nedospasov S.A.;
RT "Cloning and characterization of TNKL, a member of tankyrase gene
RT family.";
RL Genes Immun. 2:52-55(2001).
RN [3]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC TISSUE=Liver;
RX MEDLINE=21264473; PubMed=11278563;
RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,
RA Eyre H.J., Sutherland G.R., Daly R.J.;
RT "Identification of a novel human tankyrase through its interaction
RT with the adaptor protein Grb14.";
RL J. Biol. Chem. 276:17172-17180(2001).
RN [4]
RP SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RC TISSUE=Placenta;
RX MEDLINE=21443728; PubMed=11454873;
RA Kaminker P.G., Kim S.-H., Taylor R.D., Zebarjadian Y., Funk W.D.,
RA Morin G.B., Yaswen P., Campisi J.;
RT "TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
RT rapid induction of cell death upon overexpression.";
RL J. Biol. Chem. 276:35891-35899(2001).
RN [5]
RP SEQUENCE FROM N.A., FUNCTION, AND INTERACTION WITH TRF1 AND
RP LNPEP/OTASE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21661461; PubMed=11802774;

```
RA   Sbodio J.I., Lodish H.F., Chi N.-W.;
RT   "Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1
RT   (telomere-repeat-binding factor 1) and IRAP (insulin-responsive
RT   aminopeptidase).";
RL   Biochem. J. 361:451-459(2002).
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Yin Y., Gelmann E.P.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RA   Tracey A.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   SEQUENCE OF 838-1151 FROM N.A.
RC   TISSUE=Placenta;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   FUNCTION, AND ADP-RIBOSYLATION.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC       subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC       activity and can modify TRF1, and thereby contribute to the
CC       regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC       nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS. Interacts with the
CC       cytoplasmic domain of LNPEP/Otase in SLC2A4/GLUT4-vesicles. Binds
CC       to the N-terminus of Grb14 and TRF1 with its ankyrin repeat
CC       region.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC       with juxtanuclear SLC2A4/GLUT4-vesicles. Also found around the
CC       pericentriolar matrix of mitotic centromeres. During interphase, a
CC       small fraction of TNKS2 is found in the nucleus, associated with
CC       TRF1.
CC   -!- TISSUE SPECIFICITY: Highly expressed in placenta, skeletal muscle,
CC       liver, brain, kidney, heart, thymus, spinal cord, lung, peripheral
CC       blood leukocytes, pancreas, lymph nodes, spleen, prostate, testis,
CC       ovary, small intestine, colon, mammary gland, breast and breast
CC       carcinoma, and in common-type meningioma. Highly expressed in
CC       fetal liver, heart and brain.
CC   -!- PTM: ADP-ribosylated (-auto).
CC   -!- SIMILARITY: Belongs to the PARP family.
CC   -!- SIMILARITY: Contains 15 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF305081; AAG25674.1; ALT_INIT.
DR   EMBL; AF264912; AAG44694.1; -.
DR   EMBL; AF329696; AAK13463.1; -.
DR   EMBL; AF342982; AAK25811.1; -.
DR   EMBL; AF309033; AAK82330.1; -.
DR   EMBL; AF438201; AAL40795.1; -.
DR   EMBL; AL359707; CAC78760.1; -.
DR   EMBL; AK023746; BAB14665.1; ALT_INIT.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:15677; TNKS2.
DR   MIM; 607128; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 18.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation.
FT   REPEAT      57    89       ANK 1.
FT   REPEAT      90   122       ANK 2.
FT   REPEAT     123   155       ANK 3.
FT   REPEAT     210   242       ANK 4.
FT   REPEAT     243   275       ANK 5.
FT   REPEAT     276   308       ANK 6.
FT   REPEAT     363   398       ANK 7.
FT   REPEAT     399   431       ANK 8.
FT   REPEAT     432   464       ANK 9.
FT   REPEAT     525   557       ANK 10.
FT   REPEAT     558   590       ANK 11.
FT   REPEAT     591   623       ANK 12.
FT   REPEAT     678   710       ANK 13.
FT   REPEAT     711   743       ANK 14.
FT   REPEAT     744   776       ANK 15.
FT   DOMAIN     873   936       SAM.
FT   DOMAIN    1023  1162       PARP.
FT   CONFLICT   331   337       KGHSLLQ -> QRPLVAA (IN REF. 1).
FT   CONFLICT   357   361       NFKHP -> IQAS (IN REF. 1).
FT   CONFLICT   966   966       Q -> P (IN REF. 8).
SQ   SEQUENCE   1166 AA;  126917 MW;  4C8B3B8D97CEF704 CRC64;

  Query Match             99.9%;  Score 5766;  DB 1;  Length 1166;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
                 ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166
```

RESULT 2
TNK1_HUMAN
ID   TNK1_HUMAN   STANDARD;   PRT; 1327 AA.
AC   O95271; O95272;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase I) (TNKS-1) (TRF1-
DE   interacting ankyrin-related ADP-ribose polymerase).
GN   TNKS OR TNKS1 OR TIN1 OR TINF1 OR PARPL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Testis;
RX   MEDLINE=99040105; PubMed=9822378;
RA   Smith S., Giriat I., Schmitt A., de Lange T.;
RT   "Tankyrase, a poly(ADP-ribose) polymerase at human telomeres.";
RL   Science 282:1484-1487(1998).
RN   [2]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=99454782; PubMed=10523501;
RA   Smith S., de Lange T.;
RT   "Cell cycle dependent localization of the telomeric PARP, tankyrase,
RT   to nuclear pore complexes and centrosomes.";
RL   J. Cell Sci. 112:3649-3656(1999).
RN   [3]
RP   FUNCTION, AND PHOSPHORYLATION.
RX   MEDLINE=20556282; PubMed=10988299;
RA   Chi N.-W., Lodish H.F.;
RT   "Tankyrase is a Golgi-associated mitogen-activated protein kinase
RT   substrate that interacts with IRAP in GLUT4 vesicles.";
RL   J. Biol. Chem. 275:38437-38444(2000).
RN   [4]
RP   FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
RX   MEDLINE=21602874; PubMed=11739745;
RA   Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT   "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT   at human telomeres.";
RL   Mol. Cell. Biol. 22:332-342(2002).
CC   -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC   subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC   activity and can modify TRF1, and thereby contribute to the
CC   regulation of telomere length.
CC   -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =
CC   nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.
CC   -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
CC   the cytoplasmic domain of LNPEP/Otase in SLC2A4/GLUT4-vesicles.
CC   Binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC   with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is
CC   also found at nuclear pore complexes and around the pericentriolar
CC   matrix of mitotic centromeres. During interphase, a small fraction
CC   of TNKS is found in the nucleus, associated with TRF1.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=1;
CC   IsoId=O95271-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=O95271-2; Sequence=VSP_004538, VSP_004539;
CC   Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC   -!- PTM: Upon insulin-stimulation, phosphorylated on serine residues
CC   by MAPK kinases.
CC   -!- PTM: ADP-ribosylated (-auto).
CC   -!- SIMILARITY: Belongs to the PARP family.
CC   -!- SIMILARITY: Contains 15 ANK repeats.
CC   -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AF082556; AAC79841.1; -.
DR   EMBL; AF082557; AAC79842.1; -.
DR   EMBL; AF082558; AAC79843.1; -.
DR   EMBL; AF082559; AAC79844.1; -.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:11941; TNKS.
DR   MIM; 603303; -.
DR   GO; GO:0000781; C:chromosome, telomeric region; IDA.
DR   GO; GO:0003950; F:NAD ADP-ribosyltransferase activity; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0007004; P:telomerase-dependent telomere maintenance; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 19.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 17.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 15.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
KW   Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;

```
KW   Phosphorylation; Alternative splicing.
FT   REPEAT      215    247       ANK 1.
FT   REPEAT      248    280       ANK 2.
FT   REPEAT      281    313       ANK 3.
FT   REPEAT      368    400       ANK 4.
FT   REPEAT      401    433       ANK 5.
FT   REPEAT      434    466       ANK 6.
FT   REPEAT      521    556       ANK 7.
FT   REPEAT      557    589       ANK 8.
FT   REPEAT      590    622       ANK 9.
FT   REPEAT      683    715       ANK 10.
FT   REPEAT      716    748       ANK 11.
FT   REPEAT      749    781       ANK 12.
FT   REPEAT      836    868       ANK 13.
FT   REPEAT      869    901       ANK 14.
FT   REPEAT      902    934       ANK 15.
FT   DOMAIN     1030   1089       SAM.
FT   DOMAIN     1176   1327       PARP.
FT   DOMAIN        9     14       POLY-HIS.
FT   DOMAIN       27     34       POLY-PRO.
FT   DOMAIN      128    134       POLY-SER.
FT   DOMAIN      137    145       POLY-SER.
FT   VARSPLIC    641    643       EST -> GHS (in isoform 2).
FT                                /FTId=VSP_004538.
FT   VARSPLIC    644   1327       Missing (in isoform 2).
FT                                /FTId=VSP_004539.
FT   MUTAGEN    1184   1184       H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT                                WITH A-1291.
FT   MUTAGEN    1291   1291       E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
FT                                WITH A-1184.
SQ   SEQUENCE   1327 AA;  142010 MW;  E14DE985C710B957 CRC64;

  Query Match             84.8%;  Score 4889.5;  DB 1;  Length 1327;
  Best Local Similarity   83.0%;  Pred. No. 3.3e-287;
  Matches  910;  Conservative   98;  Mismatches   80;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:|||  ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||:| |||||::|||||||||||||||||||||:|| 
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||||:|||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:|||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:|||||:||| |||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||||:|||||||||||||||||||||||||:|||||||||:|||||||||||||:||
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:|||||||:|||||||||||||||||||||:|||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||:|||||||||||||||||||||||||||||||||||||||||||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: |||  |||| |||:|:|||          : :| |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              :||:  :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:|||||||||||||||||||||: |||| ||||| :   ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :|||||||||:|||||||:|||||||:::|||| |||| ||: ||:||||||||||
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              |||||||||:|||||||||||||| |||||:||||||| |||||||| |||||||||
Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095
              |||||||||||||:||
Db       1298 EQAYPEYLITYQIMKPE 1314


RESULT 3
ANK3_HUMAN
ID   ANK3_HUMAN     STANDARD;      PRT;  4377 AA.
AC   Q12955;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin 3 (ANK-3) (Ankyrin G).
GN   ANK3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain stem;
RX   MEDLINE=95138209; PubMed=7836469;
RA   Kordeli E., Lambert S., Bennett V.;
RT   "AnkyrinG. A new ankyrin gene with neural-specific isoforms localized
RT   at the axonal initial segment and node of Ranvier.";
RL   J. Biol. Chem. 270:2352-2359(1995).
CC   -!- FUNCTION: Membrane-cytoskeleton linker. The neural-specific
CC       isoforms may participate in the maintenance/targeting of ion
CC       channels and cell adhesion molecules at the nodes of Ranvier and
CC       axonal initial segments.
CC   -!- SUBUNIT: Neural-specific isoforms may be a constituent of a
CC       neurofascin/NRCAM/ankyrin G complex.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=1;
CC         Comment=A number of isoforms are produced;
CC       Name=1; Synonyms=480-kDa isoform;
CC         IsoId=Q12955-1; Sequence=Displayed;
CC   -!- TISSUE SPECIFICITY: Expressed in brain and other tissues. Isoform
```

```
CC       1 is neural-specific.
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U13616; AAA64834.1; -.
DR   PIR; A55575; A55575.
DR   HSSP; P55273; 1BI8.
DR   Genew; HGNC:494; ANK3.
DR   MIM; 600465; -.
DR   GO; GO:0006605; P:protein targeting; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT   REPEAT      73    102       ANK 1.
FT   REPEAT     106    135       ANK 2.
FT   REPEAT     139    168       ANK 3.
FT   REPEAT     172    201       ANK 4.
FT   REPEAT     203    230       ANK 5.
FT   REPEAT     234    263       ANK 6.
FT   REPEAT     267    296       ANK 7.
FT   REPEAT     300    329       ANK 8.
FT   REPEAT     333    362       ANK 9.
FT   REPEAT     366    395       ANK 10.
FT   REPEAT     399    428       ANK 11.
FT   REPEAT     432    461       ANK 12.
FT   REPEAT     465    494       ANK 13.
FT   REPEAT     498    527       ANK 14.
FT   REPEAT     531    560       ANK 15.
FT   REPEAT     564    593       ANK 16.
FT   REPEAT     597    626       ANK 17.
FT   REPEAT     630    659       ANK 18.
FT   REPEAT     663    692       ANK 19.
FT   REPEAT     696    725       ANK 20.
FT   REPEAT     729    758       ANK 21.
FT   REPEAT     762    791       ANK 22.
FT   REPEAT     795    825       ANK 23.
FT   DOMAIN    1519   1898       SER-RICH.
FT   DOMAIN    4090   4174       DEATH.
SQ   SEQUENCE   4377 AA;  480399 MW;  F42379E55768B684 CRC64;

  Query Match             15.2%;  Score 875.5;  DB 1;  Length 4377;
  Best Local Similarity   30.5%;  Pred. No. 4.2e-44;
  Matches  273;  Conservative  126;  Mismatches  362;  Indels  135;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:| |     |   || |   | |||| :|: :||: ||:    :|||
Db         85 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 144

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
                : ||  :  ::|   || :||  ::   || | | :|     |     :|  : |   :
Db        145 YMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 204

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
                |  |   ::  |  |||   : |          ||||:|| |  : :  |||   | |
Db        205 ALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 264

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |   |: : :||:  || ::|     ||||  |    :|  :||
Db        265 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRA 324

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |    |                      | | |  | : |      ||:   : :|
Db        325 A-PILSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 356

Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | || | |     ||||:| :|
Db        357 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTPLHIACKKNRI 413

Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
              |:|:::|| | : |:   | | :| ||: ||:      |: :|  ||  :::| ||| |
Db        414 KVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRGETALHMA 473

Qy        408 -----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVNCRDIEGR 459
                  | |: |:|:|  :      | |  |  :|: |  : |::|     | |     |
Db        474 ARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGASPNAATTSG- 532

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
              |||| :|      | :|| ||| :    | | ||| |  ||  ||| ||::  |
Db        533 -YTPLHLSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVANLLLQKSASP 591

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL- 577
              : |    ||| ||      ::  ||| || |    ::| ||| :  |    ||   |
Db        592 DAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATTLL 651

Qy        578 ----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNN 625
                       :| |::  ||::| :  |   |   : |||  :  |   |||||||  :
Db        652 EYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNANVNLSNKSG--LTPLHLAAQEDR 709

Qy        626 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
              : ||| |:  || |:|| | |  |||    ||:: :   |::::| ||| |  :||||:
Db        710 VNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTPLHQ 769

Qy        686 AAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLT-AAMPPSALP 738
              |||:| |  :  :|| : | |       | | |  :     :|  |   ::|  |  : :
Db        770 AAQQGHTHIINVLLQNNASPNELTVNGNTALGIARRLGYISVVDTLKIVTEETMTTTTVT 829

Qy        739 SCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 795
              :|   |: :| |                |    :  |:: : ||
Db        830 EKHKMNVPETMNEVLD-----------MSDDEVRKANAPEMLSDG-------------- 863

Qy        796 GASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ-ITLDVLVEMGHKELKEIG 850
                                          |:: |: | |  :| |    :| ::|||:|
Db        864 -------------------------EYISDVEEGEDAMTGDTDKYLGPQDLKELG 893


RESULT 4
ANK2_HUMAN
ID   ANK2_HUMAN     STANDARD;      PRT;  3924 AA.
AC   Q01484; Q01485;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN   ANK2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC   TISSUE=Brain stem;
RX   MEDLINE=91302466; PubMed=1830053;
RA   Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT   "Isolation and characterization of cDNAs encoding human brain
RT   ankyrins reveal a family of alternatively spliced genes.";
RL   J. Cell Biol. 114:241-253(1991).
RN   [2]
```

```
RP   REVISIONS.
RA   Carpenter S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain stem;
RX   MEDLINE=94075409; PubMed=8253844;
RA   Chan W., Kordeli E., Bennett V.;
RT   "440-kD ankyrinB: structure of the major developmentally regulated
RT   domain and selective localization in unmyelinated axons.";
RL   J. Cell Biol. 123:1463-1473(1993).
RN   [4]
RP   SEQUENCE OF 463-495 FROM N.A.
RX   MEDLINE=92009921; PubMed=1833308;
RA   Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
RA   Lux S.E., Ward D.C., Forget B.G.;
RT   "Isolation and chromosomal localization of a novel nonerythroid
RT   ankyrin gene.";
RL   Genomics 10:858-866(1991).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements. Also bind to cytoskeletal proteins.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=Q01484-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q01484-2; Sequence=VSP_000267, VSP_000268;
CC       Name=3;
CC         IsoId=Q01484-3; Sequence=VSP_000268;
CC   -!- TISSUE SPECIFICITY: Plasma membrane of neurons as well as glial
CC       cells throughout the brain.
CC   -!- PTM: Phosphorylated at multiple sites by different protein kinases
CC       and each phosphorylation event regulates the protein's structure
CC       and function (Potential).
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X56957; CAA40278.1; -.
DR   EMBL; X56958; CAA40279.2; -.
DR   EMBL; Z26634; CAB42644.1; -.
DR   EMBL; M37123; AAA62828.1; -.
DR   PIR; S37431; S37431.
DR   HSSP; P42771; 1DC2.
DR   Genew; HGNC:493; ANK2.
DR   MIM; 106410; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Phosphorylation.
FT   REPEAT       63     92       ANK 1.
FT   REPEAT       96    125       ANK 2.
FT   REPEAT      129    158       ANK 3.
FT   REPEAT      162    191       ANK 4.
FT   REPEAT      193    220       ANK 5.
FT   REPEAT      232    261       ANK 6.
FT   REPEAT      265    294       ANK 7.
FT   REPEAT      298    327       ANK 8.
FT   REPEAT      331    360       ANK 9.
FT   REPEAT      364    393       ANK 10.
FT   REPEAT      397    426       ANK 11.
FT   REPEAT      430    459       ANK 12.
FT   REPEAT      463    492       ANK 13.
FT   REPEAT      496    525       ANK 14.
FT   REPEAT      529    558       ANK 15.
FT   REPEAT      562    591       ANK 16.
FT   REPEAT      595    624       ANK 17.
FT   REPEAT      628    657       ANK 18.
FT   REPEAT      661    690       ANK 19.
FT   REPEAT      694    723       ANK 20.
FT   REPEAT      727    756       ANK 21.
FT   REPEAT      760    789       ANK 22.
FT   REPEAT      793    822       ANK 23.
FT   DOMAIN     1773   1950       REPEAT-RICH REGION.
FT   REPEAT     1773   1784       REPEAT A.
FT   REPEAT     1785   1796       REPEAT A.
FT   REPEAT     1797   1808       REPEAT A.
FT   REPEAT     1809   1820       REPEAT A.
FT   REPEAT     1821   1832       REPEAT A.
FT   REPEAT     1833   1844       REPEAT A.
FT   REPEAT     1845   1856       REPEAT A.
FT   REPEAT     1857   1867       REPEAT A (APPROXIMATE).
FT   REPEAT     1868   1879       REPEAT A.
FT   REPEAT     1880   1891       REPEAT A.
FT   REPEAT     1892   1902       REPEAT A (APPROXIMATE).
FT   REPEAT     1903   1914       REPEAT A.
FT   REPEAT     1915   1926       REPEAT A.
FT   REPEAT     1927   1938       REPEAT A.
FT   REPEAT     1939   1950       REPEAT A.
FT   DOMAIN     3536   3620       DEATH.
FT   VARSPLIC   1039   1039       Q -> QFLGKLHLPTAPPPLNEGESLVSRILQLGPPGTK
FT                                (in isoform 2).
FT                                /FTId=VSP_000267.
FT   VARSPLIC   1444   3528       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_000268.
FT   CONFLICT    475    476       GQ -> PE (IN REF. 4).
FT   CONFLICT    971    971       I -> S (IN REF. 1).
FT   CONFLICT   3581   3582       QY -> HA (IN REF. 1).
FT   CONFLICT   3586   3586       I -> Y (IN REF. 1).
SQ   SEQUENCE   3924 AA;  430337 MW;  52AC496C428E29D2 CRC64;

  Query Match           14.9%;  Score 860.5;  DB 1;  Length 3924;
  Best Local Similarity  29.8%;  Pred. No. 2.9e-43;
  Matches  283;  Conservative  127;  Mismatches  386;  Indels  153;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |   :|: ||  |:|| :    |   || |   | |||| :|:: ||: ||:   :|||
Db         75 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPL 134

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL--- 115
              : || :  |||   ||::||  :    || |    ||       | |:|     | ::
Db        135 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLP 194

Qy        116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS----------TPLHLAAGYNRVKIVQL 163
               |   |   :: |  |||      | | :| :              ||||:|| |  | :  |
Db        195 ALHIAARKDDTKSAALLLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 251

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
              ||  || |    :  : ||| |   |:  : :||:  |  ::|       |||| ||
Db        252 LLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 311

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE-ADVTRIKK 282
              :|  |||  || | |  |                          | | |  | : |     |
Db        312 DQVVELLLERGA-PLLARTKN--------------------GLSPLHMAAQGDHVECVK 349

Qy        283 HLSLEMVNFKHP-----QTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPL 337
              ||    :  | |        : |||| ||   :    :: :||| | || | :      |||
Db        350 HL----LQHKAPVDDVTLDYLTALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPL 402
```

```
Qy        339 HVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISL 398
              |:| :|    |:|::||: | : |:   | | :| ||: |||    |||  |  |:: ::
Db        403 HIACKKNRIKVMELLVKYGASIQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNI 462

Qy        399 QGFTALQMGN-----ENVQQLLQEGI----------------SLGNSEADRQLLE---- 432
              :| ||| |       | |: ||: |                 || :|  : ||:
Db        463 RGETALHMAARAGQVEVVRCLLRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAH 522

Qy        433 --------------AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                           :|: | |:    |   : :   : :  |||| || |  : | :
Db        523 PDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKL 581

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
              ||  |   :  | || ||| |  | : :|| ||:: ||  :      :|||| || | :
Db        582 LLQRRAAADSAGKNGLTPLHVAAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQ 641

Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597
              :|   || :||:     : | ||| |  ::| ||:  ||    | :  : |  |
Db        642 MQIASTLLNYGAETNIVTKQGVTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----- 696

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                          | ||||  : : ||: | :|||| :|  | |  ||  |  ||
Db        697 ----------------TSLHLAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYG 739

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717
              :| :   |:|  | |||  |  :||||:|||:| | :  :|| ||| |       | | |
Db        740 NVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALA 799

Qy        718 L------VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 771
              :      :|  |  ::|  :  :       |    : :  |    : |          :|
Db        800 IAKRLGYISVVDTLKVVTEEVTTTTTTITEK----HKLNVPETMTEVL--------DVSD 847

Qy        772 ASSLDNLSGSFSEL-----------SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RN 818
                  | ::|   |               |: ||   :| : |       |   |: |   |:
Db        848 EEGDDTMTGDGGEYLRPEDLKELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRS 907

Qy        819 LGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERL 867
                | |    :  |:   :|  | :   ::  :   |  : ::|  | | |
Db        908 HTLSHAS--YLRDSAVMDDSVVIPSHQVSTLAKEAERNSYRLSWGTENL 954
```

```
RESULT 5
ANK1_HUMAN
ID   ANK1_HUMAN     STANDARD;      PRT;  1880 AA.
AC   P16157;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN   ANK1 OR ANK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC   TISSUE=Hematopoietic;
RX   MEDLINE=90158830; PubMed=2137557;
RA   Lux S.E., John K.M., Bennett V.;
RT   "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT   structure with homology to tissue-differentiation and cell-cycle
RT   control proteins.";
RL   Nature 344:36-42(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90175370; PubMed=1689849;
RA   Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
RA   Cheung M.C., Kan Y.W., Palek J.;
RT   "cDNA sequence for human erythrocyte ankyrin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734(1990).
RN   [3]
RP   VARIANT HS ILE-462.
RX   MEDLINE=96225450; PubMed=8640229;
RA   Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
RA   Dornwell M., Herbers J., Kugler W., Oezcan R., Pekrun A.,
RA   Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
RT   "Ankyrin-1 mutations are a major cause of dominant and recessive
RT   hereditary spherocytosis.";
RL   Nat. Genet. 13:214-218(1996).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements; bind to the erythrocyte membrane protein band 4.2, to
CC       Na-K ATPase, to the lymphocyte membrane protein GP85, and to the
CC       cytoskeletal proteins fodrin, tubulin, vimentin and desmin.
CC       Erythrocyte ankyrins also link spectrin (beta chain) to the
CC       cytoplasmic domain of the erythrocytes anion exchange protein;
CC       they retain most or all of these binding functions.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA
CC       MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1; Synonyms=2.1;
CC         IsoId=P16157-1; Sequence=Displayed;
CC       Name=2; Synonyms=2.2;
CC         IsoId=P16157-2; Sequence=VSP_000264, VSP_000265;
CC       Name=3;
CC         IsoId=P16157-3; Sequence=VSP_000266;
CC   -!- PTM: Regulated by phosphorylation.
CC   -!- PTM: Palmitoylated.
CC   -!- DISEASE: Defects in ANK1 are a cause of hereditary spherocytosis
CC       (HS) [MIM:182900]. Inheritance can be autosomal dominant or
CC       recessive.
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16609; CAA34610.1; -.
DR   EMBL; M28880; AAA51732.1; -.
DR   PIR; A35049; A35049.
DR   PIR; S08275; SJHUK.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:492; ANK1.
DR   MIM; 182900; -.
DR   GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Phosphorylation; Lipoprotein; Disease mutation; Elliptocytosis;
KW   Polymorphism.
FT   INIT_MET      0      0
FT   DOMAIN        1    826       89 kDa DOMAIN
FT                                (ANION EXCHANGE PROTEIN BINDING DOMAIN).
FT   DOMAIN      827   1381       62 kDa DOMAIN (SPECTRIN BINDING DOMAIN).
FT   DOMAIN     1382   1880       55 kDa REGULATORY DOMAIN
FT                                (REGULATES THE BINDING OF ANKYRIN TO
FT                                SPECTRIN AND THE BAND 3 PROTEIN).
FT   REPEAT       43     72       ANK 1.
```

```
FT   REPEAT       76    105       ANK 2.
FT   REPEAT      109    138       ANK 3.
FT   REPEAT      142    171       ANK 4.
FT   REPEAT      173    200       ANK 5.
FT   REPEAT      204    233       ANK 6.
FT   REPEAT      237    266       ANK 7.
FT   REPEAT      270    299       ANK 8.
FT   REPEAT      303    332       ANK 9.
FT   REPEAT      336    365       ANK 10.
FT   REPEAT      369    398       ANK 11.
FT   REPEAT      402    431       ANK 12.
FT   REPEAT      435    464       ANK 13.
FT   REPEAT      468    497       ANK 14.
FT   REPEAT      501    530       ANK 15.
FT   REPEAT      534    563       ANK 16.
FT   REPEAT      567    596       ANK 17.
FT   REPEAT      600    629       ANK 18.
FT   REPEAT      633    662       ANK 19.
FT   REPEAT      666    695       ANK 20.
FT   REPEAT      699    728       ANK 21.
FT   REPEAT      732    761       ANK 22.
FT   REPEAT      765    794       ANK 23.
FT   DOMAIN     1402   1486       DEATH.
FT   VARSPLIC   1512   1873       Missing (in isoform 2).
FT                                /FTId=VSP_000264.
FT   VARSPLIC   1874   1874       H -> D (in isoform 2).
FT                                /FTId=VSP_000265.
FT   VARSPLIC   1849   1880       TVEGPLEDPSELEVDIDYFMKHSKDHTSTPNP -> ELRGS
FT                                GLQPDLIEGRKGAQIVKRASLKRGKQ (in isoform
FT                                3).
FT                                /FTId=VSP_000266.
FT   VARIANT      20     20       R -> T.
FT                                /FTId=VAR_000595.
FT   VARIANT     462    462       V -> I (in HS).
FT                                /FTId=VAR_000596.
FT   VARIANT     618    618       R -> H (in Brueggen).
FT                                /FTId=VAR_000597.
FT   VARIANT     749    749       V -> A.
FT                                /FTId=VAR_000598.
FT   VARIANT     844    844       D -> E.
FT                                /FTId=VAR_000599.
FT   VARIANT    1285   1285       E -> D.
FT                                /FTId=VAR_000601.
FT   VARIANT    1391   1391       S -> T.
FT                                /FTId=VAR_000600.
FT   VARIANT    1591   1591       D -> N (in Duesseldorf).
FT                                /FTId=VAR_000602.
FT   VARIANT    1698   1698       R -> D.
FT                                /FTId=VAR_000603.
FT   CONFLICT    229    229       A -> S (IN REF. 2).
FT   CONFLICT   1545   1545       V -> I (IN REF. 2).
SQ   SEQUENCE   1880 AA;  206145 MW;  1C5F5E7EFD1CD428 CRC64;

  Query Match             14.6%;  Score 842;  DB 1;  Length 1880;
  Best Local Similarity   25.5%;  Pred. No. 1.2e-42;
  Matches  279;  Conservative  152;  Mismatches  356;  Indels  306;  Gaps   27;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |      :|||
Db         88 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 147

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                |  :|  :|   |: :| :  :          || | |||          |||
Db        148 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 198

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
              ::|  :::|        |||::  |  :  :                 ||||:|:
Db        199 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 248

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
              | :|:|||  || :  | | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        249 GNVIMVRLLLDRGAQIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 308

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:           ||
Db        309 HMAAQGDHLDCVRLLLQYDAE----------ID-------------------------- 331

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                  | | || ||   :    :: ::|| |||  | :
Db        332 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 371

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :|  |  |:|:::|  | ::|:   | | || |:: |||   : ||  |  ||:
Db        372 TPLHIACKKNHVRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 431

Qy        396 ISLQGFTALQMG-----NENVQQLLQE----------------GISLGNSEADRQLLE- 432
              :::  | | |        |  : |||                    :|::   : |||
Db        432 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKADDQTPLHCAARIGHTNMVKLLLEN 491

Qy        433 ----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
                             ||: | |||| |   : | |  :|  |||| || | :|
Db        492 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVR 549

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
              | | ||:  |  :|  | || ||| |  : : :: :||:  |   : :  |  :|||| |
Db        550 VAELLLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 608

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
              | : : |: : |||:|     ::  | ||| |  ::|  ::  ||
Db        609 AKQNQVEVARSLLQYGGSANAESVQGVTPLHLAAQEGHAEMVALL---------------- 653

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
                  ||   | |  :  |  |||||  |   :: ||: |::||  |:|  : |  |||
Db        654 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHV 706

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
              |: ||:: :   |::: | |||  |  ::|||:|||:| | :  ||| :|| |   : :|
Db        707 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDG 766

Qy        713 QTPLDL-------------VSADDVSALLTAAMPPSALPSC------------------ 740
              |||  :             |  |: | :| :     :  |
Db        767 TTPLAIAKRLGYISVTDVLKVVTDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISF 826

Qy        741 -----------------YKPQVLNGVRSPGATADALSSGPSSPSSLS-----AASSLDN 777
                              : |::  | ||           |  | ::      :   :
Db        827 KAERRDSRDVDEEKELLDFVPKLDQVVESPAI--------PRIPCAMPETVVIRSEEQEQ 878

Qy        778 LSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDV 837
              |  : | | : ||  || : ::     |     :  |:
Db        879 ASKEYDEDSLIPSSPATETSDNISPVASPVHTGFLVSFM--------------------- 917

Qy        838 LVEMGHKELKEIGINAYGHRHKLIKGVE-------------RLISGQQGLNPYLTLNTSG 884
                      :   | :  | ||  :: |                ||:  |:   |      |
Db        918 --------VDARGGSMRGSRHNGLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEG 969

Qy        885 SGTILIDLSP--------------------DDKEFQSVEEEMQSTVREHRDGGHAGGIF 923
                : :| | |                    |:|   :  |  |  :|||
Db        970 LASRIIALGPTGAQFLSPVIVEIPHFASHGRGDRELVVLRSENGSVWKEHR--------- 1020

Qy        924 NRYNILKIQKVCN 936
              :||    : :: |
Db       1021 SRYGESYLDQILN 1033
```

```
RESULT 6
ANK1_MOUSE
ID   ANK1_MOUSE     STANDARD;      PRT;  1862 AA.
AC   Q02357;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ankyrin 1 (Erythrocyte ankyrin).
GN   ANK1 OR ANK-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=92345717; PubMed=1386265;
RA   White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;
RT   "Murine erythrocyte ankyrin cDNA: highly conserved regions of the
RT   regulatory domain.";
RL   Mamm. Genome 3:281-285(1992).
CC   -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
CC       elements; bind to the erythrocyte membrane protein band 4.2, to
CC       Na-K ATPase, to the lymphocyte membrane protein GP85, and to the
CC       cytoskeletal proteins fodrin, tubulin, vimentin and desmin.
CC       Erythrocyte ankyrins also link spectrin (beta chain) to the
CC       cytoplasmic domain of the erythrocytes anion exchange protein;
CC       they retain most or all of these binding functions.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC
CC       PLASMA MEMBRANE.
CC   -!- PTM: Regulated by phosphorylation (By similarity).
CC   -!- PTM: Acylated by palmitic acid group(s) (By similarity).
CC   -!- SIMILARITY: Contains 23 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M84756; AAA37236.1; -.
DR   PIR; I49502; I49502.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:88024; Ank1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
FT   DOMAIN        1    827       89 kDa DOMAIN (ANION EXCHANGE PROTEIN
FT                                BINDING DOMAIN).
FT   DOMAIN      828   1386       62 kDa DOMAIN (SPECTRIN BINDING
FT                                DOMAIN).
FT   DOMAIN     1387   1862       55 kDa REGULATORY DOMAIN (REGULATES
FT                                THE BINDING OF ANKYRIN TO SPECTRIN
FT                                AND THE BAND 3 PROTEIN).
FT   REPEAT       40     69       ANK 1.
FT   REPEAT       73    102       ANK 2.
FT   REPEAT      106    135       ANK 3.
FT   REPEAT      139    168       ANK 4.
FT   REPEAT      170    197       ANK 5.
FT   REPEAT      201    230       ANK 6.
FT   REPEAT      234    263       ANK 7.
FT   REPEAT      267    296       ANK 8.
FT   REPEAT      300    329       ANK 9.
FT   REPEAT      333    362       ANK 10.
FT   REPEAT      366    395       ANK 11.
FT   REPEAT      399    428       ANK 12.
FT   REPEAT      432    461       ANK 13.
FT   REPEAT      465    494       ANK 14.
FT   REPEAT      498    527       ANK 15.
FT   REPEAT      531    560       ANK 16.
FT   REPEAT      564    593       ANK 17.
FT   REPEAT      597    626       ANK 18.
FT   REPEAT      630    659       ANK 19.
FT   REPEAT      663    692       ANK 20.
FT   REPEAT      696    725       ANK 21.
FT   REPEAT      729    758       ANK 22.
FT   REPEAT      762    791       ANK 23.
FT   DOMAIN     1399   1483       DEATH.
SQ   SEQUENCE   1862 AA;  204242 MW;  AE6B85B5B29001E5 CRC64;

  Query Match             14.5%;  Score 837.5;  DB 1;  Length 1862;
  Best Local Similarity   29.5%;  Pred. No. 2.3e-42;
  Matches  237;  Conservative  127;  Mismatches  262;  Indels  177;  Gaps   20;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |   | ||| || :||: |      :|||
Db         85 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 144

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                | :|  :|   |: :| :  :         || | |||         |||
Db        145 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 195

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS----------------TPLHLAAGY 155
              ::|  :::|          |||::  |  :  :                ||||:|:
Db        196 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 245

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
                | :|:|||  || :  : | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        246 GNVIMVRLLLDRGAQIETRTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 305

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:          ||
Db        306 HMAAQGDHLDCVRLLLQYNAE----------ID------------------------- 328

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                | | || ||   :    :: ::|| ||| | :
Db        329 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 368

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : ||  |  ||:
Db        369 TPLHIACKKNHIRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 428

Qy        396 ISLQGFTALQM----GNENVQQLLQEGISLGNSEAD----------------RQLLE- 432
              :::  | | |     |:  | : | :  :  |::|                : |||
Db        429 SNVKVETPLHMAARAGHTEVAKYLLQNKAKANAKADDQTPLHCAARIGHTGMVKLLLEN 488

Qy        433 ----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
                             ||: | |:|   |  : |  |   :|   |||| || | :|
Db        489 GASPNLATTAGHTPLHTAAREGHVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVR 546

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
              : | ||:| |  :|  | || ||| |  : : :: :||:  |   :  :  |  :|||| |
Db        547 LAELLLEHDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 605

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
              | : : |: : |||:|     ::  | |||  |  ::| |::  ||
Db        606 AKQNQIEVARSLLQYGGSANAESVQGVTPLHLAAQEGHTEMVALL---------------- 650

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
                  ||   | |  :  |   |||||  :   :: ||: |::||  |:|  : |  |||
Db        651 -----LSKQANGNLGNKSG--LTPLHLVSQEGHVLVADVLIKHGVTVDATTRMGYTPLHV 703

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
              |: ||:: :   |::: | |||   |  ::|||:|||:| | :  ||| :|| |   :  |
Db        704 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNG 763

Qy        713 QTPLDL------VSADDVSALLT 729
              |||  :      :|  ||  ::|
Db        764 TTPLAIAKRLGYISVTDVLKVVT 786


RESULT 7
Y379_HUMAN
ID   Y379_HUMAN     STANDARD;      PRT;  1059 AA.
```

```
AC   O15084;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical ANK-repeat protein KIAA0379 (Fragment).
GN   KIAA0379.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=97349984; PubMed=9205841;
RA   Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. VII.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 4:141-150(1997).
CC   -!- SIMILARITY: Contains at least 27 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB002377; BAA20833.2; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 28.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 28.
DR   PROSITE; PS50088; ANK_REPEAT; 24.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT       46     75       ANK 1.
FT   REPEAT       79    108       ANK 2.
FT   REPEAT      112    141       ANK 3.
FT   REPEAT      145    174       ANK 4.
FT   REPEAT      178    207       ANK 5.
FT   REPEAT      211    240       ANK 6.
FT   REPEAT      244    273       ANK 7.
FT   REPEAT      277    307       ANK 8.
FT   REPEAT      311    340       ANK 9.
FT   REPEAT      344    373       ANK 10.
FT   REPEAT      377    406       ANK 11.
FT   REPEAT      410    439       ANK 12.
FT   REPEAT      443    472       ANK 13.
FT   REPEAT      476    506       ANK 14.
FT   REPEAT      510    540       ANK 15.
FT   REPEAT      555    584       ANK 16.
FT   REPEAT      588    617       ANK 17.
FT   REPEAT      622    651       ANK 18.
FT   REPEAT      658    687       ANK 19.
FT   REPEAT      691    720       ANK 20.
FT   REPEAT      724    753       ANK 21.
FT   REPEAT      761    790       ANK 22.
FT   REPEAT      793    823       ANK 23.
FT   REPEAT      828    857       ANK 24.
FT   REPEAT      861    891       ANK 25.
FT   REPEAT      895    924       ANK 26.
FT   REPEAT      931    960       ANK 27.
SQ   SEQUENCE   1059 AA;  113465 MW;  C1F55E6CFE494770 CRC64;

  Query Match             12.3%;  Score 707.5;  DB 1;  Length 1059;
  Best Local Similarity   27.2%;  Pred. No. 6.9e-35;
  Matches  259;  Conservative  130;  Mismatches  367;  Indels  197;  Gaps   29;
```

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         25 GDPDEVRALIFKKEDVNFQDNEKRTPLHAAAYLGDAEIIELLILSGARVNAKDSKWLTPL 84

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
Db         85 HRAVASCSEEAVQVLLKHSADVNARDKNWQTPLHIA-AANKAV---------------- 126

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPL 182
Db        127 ---KCAEALVPLLSNVNVSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFDKKDRRAI 183

Qy        183 HNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNC 242
Db        184 HWAAYMGHIEVVKLLVSHGAEVTCKDKKSYTPLHAAASSGMISVVKYLLDLGVDMNEPNA 243

Qy        243 HNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETA 301
Db        244 YGNTPLHVAC-----------YNGQDVV----------VNELIDCGAIVNQKNEKGF-TP 281

Qy        302 LHCAAASPYPKRKQIC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
Db        282 LHFAAASTH---GALCLELLVGNGADVNMKSKDGKTPLHMTALHGRFSRSQTIIQSGAVI 338

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQQLL 415
Db        339 DCEDKNGNTPLHIAARYGHELLINTLITSGADTAKRGIHGMFPLHLAALSGFSDCCRKLL 398

Qy        416 QEGISLGN-SEADRQLLEAAKA-GDVETVKKL------------------------CTVQ 449
Db        399 SSGFDIDTPDDFGRTCLHAAAAGGNLECLNLLLNTGADFNKKDKFGRSPLHYAAANCNYQ 458

Qy        450 ----------SVNCRDIEGRQSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGGLVPLH 498
Db        459 CLFALVGSGASVN--DLDERGCTPLHYAATSDTDGKCLEYLLRNDANPGIRDKQGYNAVH 516

Qy        499 NACSYGHYEVAELLVKHGAV-----VNVADLWK-------FTPLHEAAAKGKYEICKLLL 546
Db        517 YSAAYGHRLCLQLIASETPLDVLMETSGTDMLSDSDNRATISPLHLAAYHGHHQALEVLV 576

Qy        547 QHGADPTKKNRDGNTPLDLVK-DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVN 605
Db        577 QSLLDLDVRNSSGRTPLDLAAFKGHVECVDVLINQGASI-LVKDYILKR----------- 624

Qy        606 CRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNA---QDKGGLIPLHNAASYGHVDVA 662
Db        625 ---------TPIHAAATNGHSECLRLLIGNAEPQNAVDIQDGNGQTPLMLSVLNGHTDCV 675

Qy        663 ALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA- 721
Db        676 YSLLNKGANVDAKDKWGRTALHRGAVTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAAC 735

Qy        722 ---DDVSALLTAAMPPSALP------------SCYKPQVLNG--------------VRS 751
Db        736 GHIGVLGALLQSAASMDANPATADNHGYTALHWACY-----NGHETCVELLLEQEVFQKT 790

Qy        752 PGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSL----------- 800
Db        791 EGNAFSPLHCAVINDNEGAAEMLIDTLG------ASIVNATDSKGRTPLHAAAFTDHVEC 844

Qy        801 ------EKKEVPGVDFSITQFVRNLGLEHLMDIFEREQI-TLDVLVEMGHKEL 846
Db        845 LQLLLSHNAQVNSVD--------STGKTPLMMAAENGQTNTVEMLVSSASAEL 889
```

```
RESULT 8
YB23_HUMAN
ID   YB23_HUMAN     STANDARD;      PRT;   768 AA.
AC   Q9ULJ7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein KIAA1223 (Fragment).
```

```
GN   KIAA1223.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K.-I., Kikuno R., Hirosawa M., Nomura N.,
RA   Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
CC   -!- SIMILARITY: Contains at least 14 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB033049; BAA86537.1; -.
DR   HSSP; P42771; 1DC2.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 14.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   NON_TER       1      1
FT   REPEAT        1     11       ANK 1.
FT   REPEAT       15     44       ANK 2.
FT   REPEAT       48     82       ANK 3.
FT   REPEAT       86    115       ANK 4.
FT   REPEAT      119    148       ANK 5.
FT   REPEAT      152    181       ANK 6.
FT   REPEAT      185    214       ANK 7.
FT   REPEAT      218    247       ANK 8.
FT   REPEAT      251    280       ANK 9.
FT   REPEAT      284    313       ANK 10.
FT   REPEAT      317    346       ANK 11.
FT   REPEAT      350    379       ANK 12.
FT   REPEAT      383    412       ANK 13.
FT   REPEAT      416    446       ANK 14.
SQ   SEQUENCE   768 AA;  82819 MW;  2913B69BE2DFE06D CRC64;

  Query Match             9.9%;  Score 571;  DB 1;  Length 768;
  Best Local Similarity   26.1%;  Pred. No. 7.4e-27;
  Matches  207;  Conservative  104;  Mismatches  255;  Indels  228;  Gaps   20;

Qy         44 LLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAK 103
Db          2 LLQHGAEVNKADNEGRTALIAAAYMGHREIVEHLLDHGAEVNHEDVDGRTALSVA----- 56

Qy        104 AVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQL 163
Db         57 ------------------------ALCVP------ASKGHAS------------VVSL 72

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
Db         73 LIDRGAEVDHCDKDGMTPLLVAAYEGHVDVVDLLLEGGADVDHTDNNGRTPLLAAASMGH 132

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ-AAREADVTRIK 282
Db        133 ASVVNTLLFWGA-----------AVD----------SIDSEGRTVLSIASAQGNVEVVR 170

Qy        283 KHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVAS 342
Db        171 TLLDRGLDENHRDDAGWTPLHMAA---FEGHRLICEALIEQGARTNEIDNDGRIPFILAS 227

Qy        343 EKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFT 402
Db        228 QEGHYDCVQILLENKSNIDQRGYDGRNALRVAALEGHRDIVELLFSHGAD---------- 277

Qy        403 ALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQST 462
Db        278 ------------------------------------------------VNCKDADGRPT- 288

Qy        463 PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 522
Db        289 -LYILALENQLTMAEYFLENGANVEASDAEGRTALHVSCWQGHMEMVQVLIAYHADVNAA 347

Qy        523 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAA 582
Db        348 DNEKRSALQSAAWQGHVKVVQLLIEHG------------------------------A 375

Qy        583 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642
Db        376 VVD----------------------HTCNQGATALCIAAQEGHIDVVQVLLEHGADPNHA 413

Qy        643 DKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHG 702
Db        414 DQFGRTAMRVAAKNGHSQIIKLLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKV---- 467

Qy        703 ADPTLK-NQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSS 761
Db        468 QSLTIKSNSSGST-----GGGD--------MQPS----------LRGL--PNGPTHAFSS 502

Qy        762 GPSSPSSL--SAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNL 819
Db        503 PSESPDSTVDRQKSSLSNNSLKSSKNSSLRTTSSTATAQT-----VPIDSFHNLSFTEQI 557

Qy        820 GLEHLMDIFEREQI 833
Db        558 QQHSLPRSRSRQSI 571


RESULT 9
LATA_LATMA
ID   LATA_LATMA     STANDARD;      PRT;  1401 AA.
AC   P23631; Q25328;
DT   01-NOV-1991 (Rel. 20, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-latrotoxin precursor.
OS   Latrodectus mactans (Black widow spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Araneoidea; Theridiidae; Latrodectus.
OX   NCBI_TaxID=6924;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tredecimguttatus; TISSUE=Venom gland;
RX   MEDLINE=91031994; PubMed=1977615;
RA   Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V.;
RT   "Cloning and structure of cDNA encoding alpha-latrotoxin from black
RT   widow spider venom.";
RL   FEBS Lett. 270:127-131(1990).
RN   [2]
RP   PARTIAL SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=91362695; PubMed=1888339;
RA   Volkova T.M., Galkina T.G., Kudelin A.B., Nazimov I.V., Grishin E.V.;
RT   "Structure of tryptic fragments of a neurotoxin from black widow
RT   spider venom.";
RL   Bioorg. Khim. 17:437-441(1991).
CC   -!- FUNCTION: Binds to high-affinity receptors that are localized in
CC       the presynaptic plasma membrane of the nerve terminal. Binding of
CC       toxin results in calcium influx into nerve terminals, activation
CC       of synaptic vesicle exocytosis, and ATP depletion.
CC   -!- PTM: Alpha-latrotoxin may be processed at its C-terminus.
CC   -!- SIMILARITY: Contains 21 ANK repeats.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X55009; CAA38753.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 20.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 20.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Toxin; Neurotoxin; Presynaptic neurotoxin; Repeat; ANK repeat; Signal.
FT   SIGNAL        1     20
FT   CHAIN        21   1401       ALPHA-LATROTOXIN.
FT   REPEAT      490    521       ANK 1.
FT   REPEAT      525    554       ANK 2.
FT   REPEAT      559    589       ANK 3.
FT   REPEAT      593    622       ANK 4.
FT   REPEAT      626    656       ANK 5.
FT   REPEAT      660    690       ANK 6.
FT   REPEAT      695    723       ANK 7.
FT   REPEAT      729    758       ANK 8.
FT   REPEAT      762    791       ANK 9.
FT   REPEAT      795    824       ANK 10.
FT   REPEAT      828    857       ANK 11.
FT   REPEAT      862    891       ANK 12.
FT   REPEAT      895    924       ANK 13.
FT   REPEAT      928    957       ANK 14.
FT   REPEAT      971   1003       ANK 15.
FT   REPEAT     1004   1033       ANK 16.
FT   REPEAT     1035   1064       ANK 17.
FT   REPEAT     1068   1097       ANK 18.
FT   REPEAT     1101   1131       ANK 19.
FT   REPEAT     1137   1166       ANK 20.
FT   REPEAT     1170   1199       ANK 21.
SQ   SEQUENCE   1401 AA;  156856 MW;  B38A22083C142A98 CRC64;

  Query Match             8.1%;  Score 464.5;  DB 1;  Length 1401;
  Best Local Similarity   23.2%;  Pred. No. 4.9e-20;
  Matches  175;  Conservative   94;  Mismatches  274;  Indels  211;  Gaps   19;

Qy        12 LQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGK 70
             |      :   |   |  |:| |   |:| :||||:: |    |::  ::  ||||  || :|
Db       513 LAKSTELNQPDKKGYTPIHVAADSGNAGIVNLLIQRGVSINSKTYHFLQTPLHLAAQRGF 572

Qy        71 IDVCIVLLQHGAEPTI----RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEK 126
             :    |::    | |    |: || |                  |  |  | |
Db       573 VTTFQRLME---SPEININERDKDGFTP--------------------LHYAIRGGERI 608

Qy       127 MMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAKDKGDLVPLHNA 185
             : | |  :::: :|     || |||   |   :   ||     |::| |: ::  || |
Db       609 LEAFLNQISIDVNAKSNTGLTPFHLAIIKNDWPVASTLLGSKKVDINAVDENNITALHYA 668

Qy       186 CSYGHYEVTE-------------------------LLVKH-----------GACVNAMDL 209
               |: | |:                         :| ||                ||  |
Db       669 AILGYLETTKQLINLKEINANVVSSPGLLSALHYAILYKHDDVASFLMRSSNVNVNLKAL 728

Qy       210 WQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSL 269
               |||| |  :  | :: ||:   |
Db       729 GGITPLHLAVIQGRKQILSLMFDIG---------------------------------- 753

Qy       270 LQAAREADVTRIKKHLSLEMVNFKHPQTHE--TALHCAAASPYPKRKQICELLLRKGANI 327
                                 || :  :| |  | || || | ||   :: ::|| :|:|
Db       754 -------------------VNIEQ-KTDEKYTPLHLAAMSKYP---ELIQILLDQGSNF 789

Qy       328 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 387
               ||    ||||:|: |  :    ::: :|      |  ||  :| ||  | |   : ::
```

```
Db       790 EAKTNSGATPLHLATFKGKSQAALILLNNEVNWRDTDENGQMPIHGAAMTGLLDVAQAII 849

Qy       388 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 447
             |   |  ::
Db       850 S--IDATVV---------------------------------------------------- 856

Qy       448 VQSVNCRDIEGRQS-TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 506
                 ||| :  | |||: ||  :  : |::| :  |||:: ::| || ||       |:
Db       857 -------DIEDKNSDTPLNLAAQNSHIDVIKYFIDQGADINTRNKKGLAPLLAFSKKGNL 909

Qy       507 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 566
             :: : |    | | :||       : |   |   | |  :           | | |
Db       910 DMVKYLFDKNANVYIADNDGMNFFYYAVQNGHLNIVKYAMSEKDKFEWSNTDNN------ 963

Qy       567 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHST--PLHLAAGYN 624
             :  :   ::       |: || :   :  ||             | |  :   ||| || |
Db       964 RRDECPNEECAISHFAVCDAVQFDRIEIVKYFVG--------TLGNFAICGPLHQAARYG 1015

Qy       625 NLEVAEYLLQH---GADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFT 681
             :|:: :||::      | :  |      ||  |:  ||  |   |:   | ||       |
Db      1016 HLDIVKYLVEEEFLSVDGSKTD----TPLCYASENGHFTVVQYLVSNGAKVNHDCGNGMT 1071

Qy       682 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 715
               : :| |   |:   | |:| |   ||  | ||
Db      1072 AIDKAITKNHLQVVQFLAANGVDFRRKNSRGTTP 1105
```

```
RESULT 10
AFY1_MOUSE
ID   AFY1_MOUSE     STANDARD;      PRT;  1169 AA.
AC   Q810B6; O54807; Q80TG6; Q80UH8;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin repeats- and FYVE domain-containing protein 1 (Ankyrin repeats
DE   hooked to a zinc finger motif).
GN   ANKFY1 OR ANKHZN OR KIAA1255.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
RX   MEDLINE=99194578; PubMed=10092534;
RA   Ito K., Ishii N., Miyashita A., Tominaga K., Kuriyama H., Maruyama H.,
RA   Shirai M., Naito M., Arakawa M., Kuwano R.;
RT   "Molecular cloning of a novel 130-kDa cytoplasmic protein, Ankhzn,
RT   containing ankyrin repeats hooked to a zinc finger motif.";
RL   Biochem. Biophys. Res. Commun. 257:206-213(1999).
RN   [2]
RP   REVISIONS, AND SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA   Maruyama H., Kuriyama H., Ishii N., Ito K., Odani S., Kuwano R.;
RT   "Genomic organization, alternative splicing, and promoter assay of the
RT   mouse Ankhzn gene.";
RL   Submitted (DEC-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 994-1169 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22579291; PubMed=12693553;
RA   Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA   Nakajima D., Nagase T., Ohara O., Koga H.;
RT   "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT   II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT   cDNAs identified by screening of terminal sequences of cDNA clones
RT   randomly sampled from size-fractionated libraries.";
RL   DNA Res. 10:35-48(2003).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Also associated with endosomes.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q810B6-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q810B6-2; Sequence=VSP_007917, VSP_007918;
```

```
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC   -!- SIMILARITY: Contains 21 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB011370; BAA24980.2; -.
DR   EMBL; AB098329; BAC67211.1; -.
DR   EMBL; AB098157; BAC67389.1; -.
DR   EMBL; AB098135; BAC67389.1; JOINED.
DR   EMBL; AB098136; BAC67389.1; JOINED.
DR   EMBL; AB098137; BAC67389.1; JOINED.
DR   EMBL; AB098138; BAC67389.1; JOINED.
DR   EMBL; AB098139; BAC67389.1; JOINED.
DR   EMBL; AB098140; BAC67389.1; JOINED.
DR   EMBL; AB098141; BAC67389.1; JOINED.
DR   EMBL; AB098142; BAC67389.1; JOINED.
DR   EMBL; AB098143; BAC67389.1; JOINED.
DR   EMBL; AB098144; BAC67389.1; JOINED.
DR   EMBL; AB098145; BAC67389.1; JOINED.
DR   EMBL; AB098146; BAC67389.1; JOINED.
DR   EMBL; AB098147; BAC67389.1; JOINED.
DR   EMBL; AB098148; BAC67389.1; JOINED.
DR   EMBL; AB098149; BAC67389.1; JOINED.
DR   EMBL; AB098150; BAC67389.1; JOINED.
DR   EMBL; AB098151; BAC67389.1; JOINED.
DR   EMBL; AB098152; BAC67389.1; JOINED.
DR   EMBL; AB098153; BAC67389.1; JOINED.
DR   EMBL; AB098154; BAC67389.1; JOINED.
DR   EMBL; AB098155; BAC67389.1; JOINED.
DR   EMBL; AB098156; BAC67389.1; JOINED.
DR   EMBL; AB098143; BAC67388.1; -.
DR   EMBL; AB098135; BAC67388.1; JOINED.
DR   EMBL; AB098136; BAC67388.1; JOINED.
DR   EMBL; AB098137; BAC67388.1; JOINED.
DR   EMBL; AB098138; BAC67388.1; JOINED.
DR   EMBL; AB098139; BAC67388.1; JOINED.
DR   EMBL; AB098140; BAC67388.1; JOINED.
DR   EMBL; AB098141; BAC67388.1; JOINED.
DR   EMBL; AB098142; BAC67388.1; JOINED.
DR   EMBL; AK122479; BAC65761.1; -.
DR   MGD; MGI:1337008; Ankfy1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF00023; ank; 21.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF01363; FYVE; 1.
DR   SMART; SM00248; ANK; 20.
DR   SMART; SM00225; BTB; 1.
DR   SMART; SM00064; FYVE; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
KW   ANK repeat; Repeat; Zinc-finger; Alternative splicing.
FT   DOMAIN       68    130       BTB.
FT   REPEAT      217    247       ANK 1.
FT   REPEAT      255    284       ANK 2.
FT   REPEAT      288    317       ANK 3.
FT   REPEAT      322    362       ANK 4.
FT   REPEAT      366    396       ANK 5.
FT   REPEAT      490    519       ANK 6.
FT   REPEAT      542    572       ANK 7.
FT   REPEAT      588    617       ANK 8.
FT   REPEAT      621    650       ANK 9.
FT   REPEAT      654    683       ANK 10.
FT   REPEAT      687    716       ANK 11.
FT   REPEAT      724    753       ANK 12.
FT   REPEAT      769    798       ANK 13.
FT   REPEAT      802    832       ANK 14.
FT   REPEAT      836    865       ANK 15.
FT   REPEAT      870    899       ANK 16.
FT   REPEAT      905    934       ANK 17.
FT   REPEAT      938    967       ANK 18.
FT   REPEAT      971   1001       ANK 19.
FT   REPEAT     1005   1037       ANK 20.
FT   REPEAT     1043   1072       ANK 21.
FT   ZN_FING    1104   1164       FYVE-TYPE.
FT   VARSPLIC    392    439       LDLELKDHEGSTALWLAVQYITVSSDQSVNPPEDLPVVNGT
FT                                SFDENSL -> YVGRPGEMQSGCEGGIIRFRAERSRRQHCT
FT                                LAGRPVHHCVFRPVCKPL (in isoform 2).
FT                                /FTId=VSP_007917.
FT   VARSPLIC    440   1169       Missing (in isoform 2).
FT                                /FTId=VSP_007918.
FT   CONFLICT     439    439       L -> F (IN REF. 1; BAC67389).
FT   CONFLICT     448    448       S -> T (IN REF. 1; BAC67389).
FT   CONFLICT     502    502       G -> V (IN REF. 1; BAC67389).
FT   CONFLICT     527    527       L -> V (IN REF. 1; BAC67389).
FT   CONFLICT     840    840       P -> S (IN REF. 1; BAC67389).
SQ   SEQUENCE   1169 AA;  128603 MW;  A78E5406A0AD0165 CRC64;

  Query Match             7.8%;  Score 449.5;  DB 1;  Length 1169;
  Best Local Similarity   24.7%;  Pred. No. 3e-19;
  Matches  199;  Conservative  119;  Mismatches  302;  Indels  187;  Gaps   26;

Qy          1 GFGRKDVV--EYLLQNGASVQARDDGGL-IPLHNACSFG-----------HAEVVNLLLR 46
              |  | |:    :|::||| | |   |    |||    :             |::   ||:
Db        296 GIQRGDLFASTFLIKNGALVNAATAGAQETPLHLVALYSPKKYSADVMSEMAQIAEALLQ 355

Qy         47 HGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQ-HGAEPTIRNTDGRTALDLA------- 98
              ||:|| :|:   ||||: :       |   |||    :  ::: :| ||| ||
Db        356 AGANPNMQDSKGRTPLHLSIMARNDCVFSQLLQCKQLDLELKDHEGSTALWLAVQYITVS 415

Qy         99 -DPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRK-STPLHLAAGYN 156
               | |           :|  : :  | :|  : | |       | | |    :  | |||
Db        416 SDQSVNPF------EDLPVVNGTSFDENSLAARLIQRGSNTDAPDVMTGNCLLQRAAGAG 469

Qy        157 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVN----------- 205
                  |   ||  : ::|    |||  || :|   :|  |:: ||  |
Db        470 NEAAALFLATSGAHANHRNKWGETPLHTACRHGLANLTAELLQQGANPNLQTEEALTLPK 529

Qy        206 --------AMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLK 257
                     |  ::  |||| | : |  :| |::|   |:       |  : : :  ||
Db        530 ESPVLMSSADSIYLQTPLHMAIAYNHPDVVSVILEQKANA----LHATNNLQIIPDFSLK 585

Qy        258 ERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQIC 317
              :                 :| :  | | :          | :|  ||
Db        586 D-------------------SRDQTVLGLAL---------WTGMHTIAAQ--------- 607

Qy        318 ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRA--- 374
              ||  ||:||:   :  | ||:| ::  :     :::|:| :|     |:|:|  |
Db        608 --LLGSGASINDTMSDGQTLLHMAIQRQDSKSALFLLEHQADINVRTQDGETALQLAIKH 665

Qy        375 -------AYCGHLQTCRLLLSYGCDPNIISLQG----FTALQMGNENV-QQLLQEGISLG 422
                    | |          : | | ::   :|    : ||     |::    |:: |
Db        666 QLPLVVDAIC----------TRGADMSVPDEKGNPPLWLALASNLEDIASTLVRHG---- 711

Qy        423 NSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 482
              |       :| ::|:                        || |   |  |   :|::
Db        712 ---CDATCWGPGPSGCLQTL--------------------LHRAVDENNESTACFLIRS 747

Qy        483 GADVHA------------KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 530
              | ||::            : : |  ||| | |:|   |   : |:: || ||  |    ||:
Db        748 GCDVNSPRQPGTNGEGEEEARDGQTPLHLAASWGLEETVQCLLEFGANVNAQDAEGRTPV 807

Qy        531 HEAAAKGKYEICKLLLQH-GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKK 589
              | | :     | :||: |   : : ::| | ||                    |:     |
Db        808 HVAISNQHSVIIQLLISHPNIELSVRDRQGLTPF---------------ACAMTYKNNK 851
```

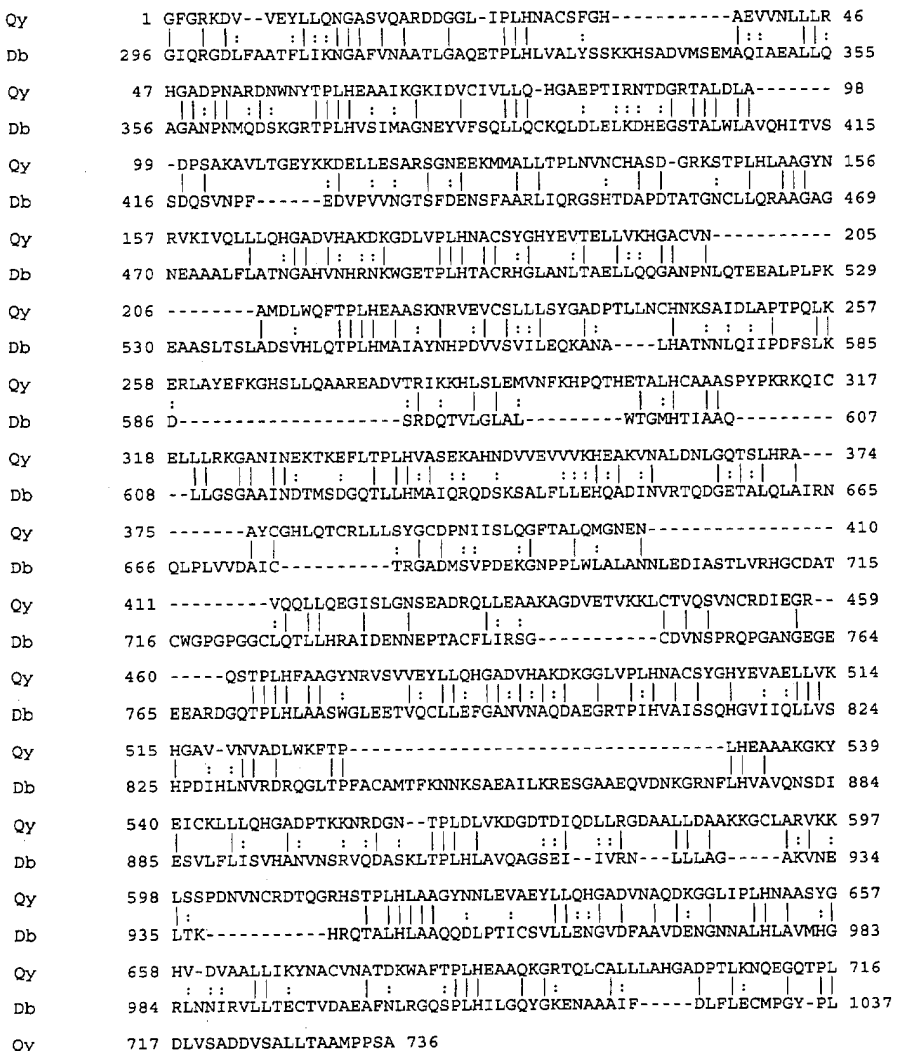
```
Qy        590 GCLARVKKLS-SPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNA--QDKGG 646
              | :|: | : : |   | :||:   ||:|   :::|   :|:   |:||:  ||
Db        852 AAEAILKRESGAAEQV---DNKGRNF--LHVAVQNSDIESVLFLISVQANVNSRVQDASK 906

Qy        647 LIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPT 706
              | ||| |   |   :   |:   | ||   |   | || |||:    :|::|| :| |
Db        907 LTPLHLAVQAGSEIIVRNLLLAGAKVNELTKHRQTALHLAAQQDLPTICSVLLENGVDFA 966

Qy        707 LKNQEGQTPLDLV----SADDVSALLT 729
                :: |   | |        ::: ||||
Db        967 AVDENGNNALHLAVMHGRLNNIRALLT 993

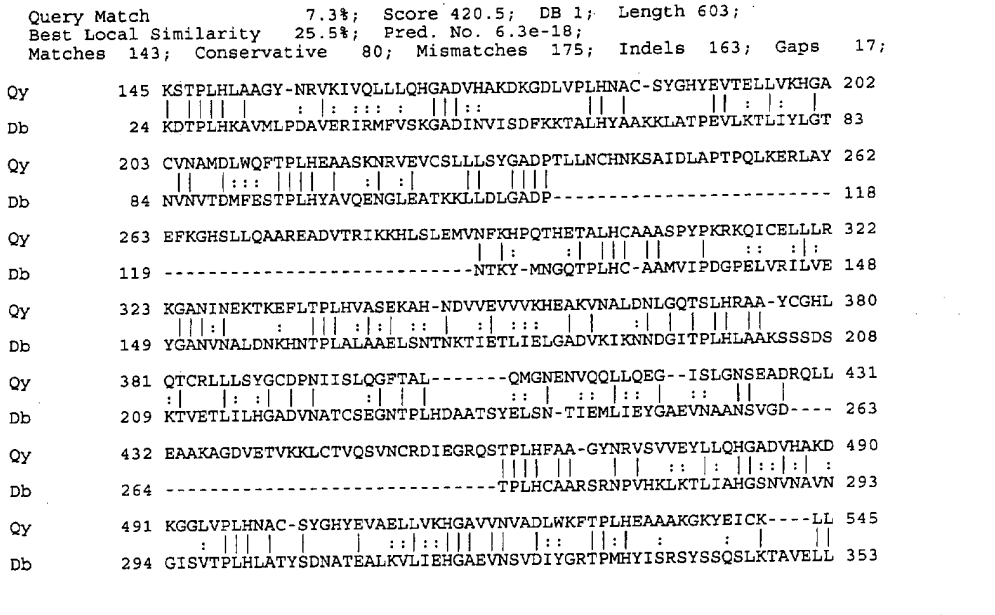

RESULT 11
AFY1_HUMAN
ID   AFY1_HUMAN     STANDARD;      PRT;  1169 AA.
AC   Q9P2R3; Q9ULG5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeats- and FYVE domain-containing protein 1 (Ankyrin repeats
DE   hooked to a zinc finger motif).
GN   ANKFY1 OR ANKHZN OR KIAA1255.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC   TISSUE=Fetal brain;
RX   MEDLINE=20400338; PubMed=10940552;
RA   Kuriyama H., Asakawa S., Minoshima S., Maruyama H., Ishii N., Ito K.,
RA   Gejyo F., Arakawa M., Shimizu N., Kuwano R.;
RT   "Characterization and chromosomal mapping of a novel human gene,
RT   ANKHZN.";
RL   Gene 253:151-160(2000).
RN   [2]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Brain;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K.-I., Kikuno R., Hirosawa M., Nomura N.,
RA   Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Also associated with endosomes.
CC   -!- TISSUE SPECIFICITY: High expression in whole adult brain and
CC       intermediate expression in all other tissues and specific brain
CC       regions examined, including fetal brain.
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC   -!- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC   -!- SIMILARITY: Contains 21 ANK repeats.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB037360; BAA90300.1; -.
DR   EMBL; AB033081; BAA86569.2; ALT_INIT.
DR   HSSP; P42773; 1BU9.
DR   Genew; HGNC:20763; ANKFY1.
DR   MIM; 607927; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR000306; Znf_FYVE.
DR   Pfam; PF00023; ank; 20.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF01363; FYVE; 1.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00225; BTB; 1.
DR   SMART; SM00064; FYVE; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 12.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS50178; ZF_FYVE; 1.
KW   ANK repeat; Repeat; Zinc-finger.
FT   DOMAIN       68    130       BTB.
FT   REPEAT      217    247       ANK 1.
FT   REPEAT      255    284       ANK 2.
FT   REPEAT      288    317       ANK 3.
FT   REPEAT      322    362       ANK 4.
FT   REPEAT      366    396       ANK 5.
FT   REPEAT      490    519       ANK 6.
FT   REPEAT      542    572       ANK 7.
FT   REPEAT      588    617       ANK 8.
FT   REPEAT      621    650       ANK 9.
FT   REPEAT      654    683       ANK 10.
FT   REPEAT      687    716       ANK 11.
FT   REPEAT      724    753       ANK 12.
FT   REPEAT      769    798       ANK 13.
FT   REPEAT      802    832       ANK 14.
FT   REPEAT      836    865       ANK 15.
FT   REPEAT      870    899       ANK 16.
FT   REPEAT      905    934       ANK 17.
FT   REPEAT      938    967       ANK 18.
FT   REPEAT      971   1001       ANK 19.
FT   REPEAT     1005   1037       ANK 20.
FT   REPEAT     1043   1072       ANK 21.
FT   ZN_FING    1104   1164       FYVE-TYPE.
FT   CONFLICT      1      4       MAEE -> AGGK (IN REF. 2).
FT   CONFLICT     48     48       E -> K (IN REF. 1).
FT   CONFLICT    115    115       M -> V (IN REF. 1).
FT   CONFLICT    207    207       A -> P (IN REF. 1).
FT   CONFLICT    235    235       F -> S (IN REF. 1).
FT   CONFLICT    270    270       S -> N (IN REF. 1).
FT   CONFLICT    283    291       DMVDKSGWS -> AWWPRVLE (IN REF. 1).
FT   CONFLICT    299    300       RG -> E (IN REF. 1).
FT   CONFLICT    322    322       A -> C (IN REF. 1).
FT   CONFLICT    336    337       KK -> RN (IN REF. 1).
FT   CONFLICT    379    379       N -> D (IN REF. 1).
FT   CONFLICT    385    385       Q -> H (IN REF. 1).
FT   CONFLICT    645    645       Q -> R (IN REF. 1).
FT   CONFLICT    652    654       TQD -> PQA (IN REF. 1).
FT   CONFLICT    687    687       K -> E (IN REF. 1).
FT   CONFLICT    713    713       D -> G (IN REF. 1).
FT   CONFLICT    784    784       MISSING (IN REF. 1).
FT   CONFLICT    795    795       N -> D (IN REF. 1).
FT   CONFLICT    798    798       A -> P (IN REF. 1).
FT   CONFLICT    807    807       I -> C (IN REF. 1).
FT   CONFLICT    827    829       DIH -> ISS (IN REF. 1).
FT   CONFLICT    835    835       R -> K (IN REF. 1).
FT   CONFLICT    849    849       N -> D (IN REF. 1).
FT   CONFLICT    860    860       E -> G (IN REF. 1).
FT   CONFLICT    874    874       F -> S (IN REF. 1).
FT   CONFLICT    889    889       F -> S (IN REF. 1).
FT   CONFLICT    901    901       V -> A (IN REF. 1).
FT   CONFLICT    905    905       S -> P (IN REF. 1).
FT   CONFLICT    913    913       A -> V (IN REF. 1).
FT   CONFLICT    916    916       A -> E (IN REF. 1).
FT   CONFLICT    933    933       N -> T (IN REF. 1).
FT   CONFLICT    940    940       Q -> K (IN REF. 1).
FT   CONFLICT    949    949       Q -> E (IN REF. 1).
FT   CONFLICT    962    962       G -> A (IN REF. 1).
FT   CONFLICT   1120   1120       T -> A (IN REF. 1).
FT   CONFLICT   1157   1157       N -> T (IN REF. 1).
SQ   SEQUENCE   1169 AA;  128398 MW;  589297CA4ACDDB56 CRC64;

  Query Match             7.6%;  Score 439;  DB 1;  Length 1169;
  Best Local Similarity   24.2%;  Pred. No. 1.3e-18;
  Matches  208;  Conservative  107;  Mismatches  323;  Indels  222;  Gaps   27;
```



```
RESULT 12
V162_FOWPV
ID   V162_FOWPV     STANDARD;      PRT;   603 AA.
AC   Q9J569;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV162.
GN   FPV162.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 14 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44506.1; -.
DR   HSSP; P25963; 1IKN.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   SMART; SM00248; ANK; 13.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       23     53       ANK 1.
FT   REPEAT       57     87       ANK 2.
FT   REPEAT       91    120       ANK 3.
FT   REPEAT      124    155       ANK 4.
FT   REPEAT      159    189       ANK 5.
FT   REPEAT      193    223       ANK 6.
FT   REPEAT      227    257       ANK 7.
FT   REPEAT      261    291       ANK 8.
FT   REPEAT      295    325       ANK 9.
FT   REPEAT      329    362       ANK 10.
FT   REPEAT      366    397       ANK 11.
FT   REPEAT      401    428       ANK 12.
FT   REPEAT      432    467       ANK 13.
FT   REPEAT      504    533       ANK 14.
SQ   SEQUENCE   603 AA;  66542 MW;  8FFE3E464DB9D16D CRC64;

  Query Match             7.3%;  Score 420.5;  DB 1;  Length 603;
  Best Local Similarity   25.5%;  Pred. No. 6.3e-18;
  Matches  143;  Conservative   80;  Mismatches  175;  Indels  163;  Gaps   17;
```



```
Qy        546 LQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVN 605
              ::||||   ||  | |||                                        |: :|:
Db        354 VEHGADIEAKNVIGGTPLS---------------------------------SACNNIE 379

Qy        606 CRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALL 665
                              :| : |  :::|||:| :|     ||::|  |   ::  ||
Db        380 -----------------YDLRLIECFIEYGADINTRDIRDETPLYSAIKY--PEIVNLL 419

Qy        666 IKYNACVNATDKWAFTPLHEA 686
              : |:|  | |:|   |||  |
Db        420 MNYSASTNITNKSNITPLESA 440


RESULT 13
V222_FOWPV
ID   V222_FOWPV     STANDARD;      PRT;   747 AA.
AC   Q9J513;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative ankyrin-repeat protein FPV222.
GN   FPV222.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 14 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44566.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001810; F-box.
DR   Pfam; PF00023; ank; 12.
DR   Pfam; PF00646; F-box; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       38     67       ANK 1.
FT   REPEAT      103    132       ANK 2.
FT   REPEAT      136    165       ANK 3.
FT   REPEAT      169    198       ANK 4.
FT   REPEAT      202    231       ANK 5.
FT   REPEAT      234    263       ANK 6.
FT   REPEAT      294    323       ANK 7.
FT   REPEAT      328    357       ANK 8.
FT   REPEAT      361    393       ANK 9.
FT   REPEAT      397    426       ANK 10.
FT   REPEAT      430    460       ANK 11.
FT   REPEAT      464    493       ANK 12.
FT   REPEAT      495    524       ANK 13.
FT   REPEAT      529    559       ANK 14.
SQ   SEQUENCE   747 AA;  85303 MW;  55F90AF2855C3D28 CRC64;

  Query Match             6.8%;  Score 394.5;  DB 1;  Length 747;
  Best Local Similarity   22.6%;  Pred. No. 3.2e-16;
  Matches  183;  Conservative  127;  Mismatches  306;  Indels  193;  Gaps   28;
```

```
Qy          4 RKDVVEY--LLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 61
              : |||      |: | :   ::      |: |    : ::: ||| ||||||   : : ||
Db         16 KNDVVSVRKYLEKGINPNEKNKDNCTMLYTAVEHRYIDIIKLLLDHGADPNIYSSDHMTP 75

Qy         62 LHEAAIKGKI-DVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||  ::    |  :  :: ::|   |: |       :                 | ::  || |
Db         76 LHSVSVIIPIRKISKLITKYGNLVTLANYRNTPFV-------------YNENRNLEIA- 120

Qy        121 SGNEEKMM----ALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDK 176
                ||:    ||:   |:       :  ||||:|:     |:|:||| |||: :|
Db        121 -----KMLIQNGALVNMNNM-------KNITPLHIASSSGSYKMVELLLLHGANTNALTS 168

Qy        177 GDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGAD 236
                || : |     ::|||:::|  ||  :    | | |     :::  |||  |||
Db        169 YGETSLHYSVSSNDLNISELLIENGTNVNVANKDSITALIIAVEIMSIDLVRLLLDKGAD 228

Qy        237 PTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFK--- 293
              :   :  : : | |   | |       |          :|| | :::: |: ::
Db        229 TNAIGL-ERFKLYVTETKQNNNILKY------LNTNNVNTNVTMINEYIASELYDWNRNS 281

Qy        294 ----------HPQTHETALHCAAASPYPKRKQICELLLRKGAN--INEKTKEFLTPLHV 340
                       | |     |      |      |:: |:||  | |  | |||      :|
Db        282 ATSKLMFRSCFEPCTVPVTLATRKGS-----KELLEILLEYGCNPDICEKTTSTYA-MHY 335

Qy        341 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 400
              |  : | ::: ::::::|  :  |     |  | |       ::|: |
Db        336 AVIRKHYEMLNILIRYDAYTDVKDRQQNTPAHYAVKLPISESCKYL-------------- 381

Qy        401 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 460
                                    :||: | |                | |  : :||
Db        382 ---------------------------KLLKLAGA------------SFNLTNRKGR- 399

Qy        461 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAV-V 519
              |||| |  ||     |:||:: | | :  |      ||: | |   :  ::|:: | |
Db        400 -TPLHTACKYNNTEAVKYLIESGCDTNIVDVMSFTPLNYAVYYEREDTVKILLESGCVDP 458

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNR-DGNTPLDLVKDGDTDIQDLLR 578
              |: |  : :|: :|  :    | |:||  | |    |  |  | : :      :|||
Db        459 NLCDYKEVSPIIQAIKRNNKNIIKMLLNAGIDIKPINECYGLHMLAALHN-----KDLL- 512

Qy        579 GDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPL-HLAAGYNNLEVAEYLLQHGA 637
                  |  | : :|     ||  |    |  ||    |  ::: : | |:: |
Db        513 ----------KWLLCTISEL----EVNGVDD---HYVPLASYVAELSDIRIMEILIEKGL 555

Qy        638 DVN---AQDKGGLIPLHNAASYGHVDVAALLI------------KYNACVNATDKWAFT 681
              |:|     |:     : :| |      :  |||            | |  :  |||:
Db        556 DLNKVTGPDETMFTMIFSATSDLRKSIIDLLISQIAADEBFSEGFKINKNMIQTDKYLLR 615

Qy        682 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS---ADDVSALLTAAMPPSAL- 737
                ||    : |:  :           :|  :| | :|:     :  |: |   ||  | :
Db        616 VYHEC----KNQVSKM-------GEIKLGDGFTMIDIYKNRRSIHVNFLARYAMQLSTID 664

Qy        738 ----------------PSCYKPQVLNGVR 750
                           |:  : ::||  :
Db        665 LREVPIYRKYLEILINPAIKRHKILNAAK 693


RESULT 14
ANR3_HUMAN
ID   ANR3_HUMAN     STANDARD;      PRT;   832 AA.
AC   P57078; Q96KH0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat
DE   domain protein 3) (PKC-delta-interacting protein kinase).
GN   ANKRD3 OR DIK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
```

```
RC   TISSUE=Fetal kidney, and Fetal lung;
RA   Shimizu N., Kudoh J., Shibuya K.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P57078-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P57078-2; Sequence=VSP_004862;
CC   -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB047783; BAB56136.1; -.
DR   EMBL; AP001743; BAA95526.1; -.
DR   HSSP; P25963; 1IKN.
DR   Genew; HGNC:496; ANKRD3.
DR   MIM; 605706; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00023; ank; 10.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR00109; TYRKINASE.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
KW   Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW   ANK repeat; Alternative splicing.
FT   DOMAIN       22    286       PROTEIN KINASE.
FT   REPEAT      485    514       ANK 1.
FT   REPEAT      518    547       ANK 2.
FT   REPEAT      551    580       ANK 3.
FT   REPEAT      584    613       ANK 4.
FT   REPEAT      617    647       ANK 5.
FT   REPEAT      651    680       ANK 6.
FT   REPEAT      684    713       ANK 7.
FT   REPEAT      717    746       ANK 8.
FT   REPEAT      750    780       ANK 9.
FT   REPEAT      782    811       ANK 10.
FT   NP_BIND      28     36       ATP (BY SIMILARITY).
FT   BINDING      51     51       ATP (BY SIMILARITY).
FT   ACT_SITE    143    143       BY SIMILARITY.
FT   VARSPLIC    278    325       Missing (in isoform 2).
FT                                /FTId=VSP_004862.
FT   CONFLICT    714    714       M -> V (IN REF. 1).
SQ   SEQUENCE   832 AA;  91610 MW;  5D8FFFD5F04F7ECB CRC64;

  Query Match             6.8%;  Score 390.5;  DB 1;  Length 832;
  Best Local Similarity   28.1%;  Pred. No. 6.6e-16;
  Matches  141;  Conservative   58;  Mismatches  176;  Indels  127;  Gaps   13;

Qy        228 SLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL 287
              || ||:  :|          |  ||  |   |::          |: |    | ::: | |
Db        434 SLSLSFEREP--------STSDLGTTDVQKKK---------LVDAIVSGDTSKLMKIL-- 474

Qy        288 EMVNFKHPQTHETALHCAAA----SPYPKRKQICELLLRKGANINEKTKEFLTPLHVASE 343
                  ||  : ||  |:   :    ::  : ||   || |   :   ||||:| |
Db        475 ------QPQDVDLALDSGASLLHLAVEAGQEECAKWLLLNNANPNLSNRRGSTPLHMAVE 528

Qy        344 KAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFT- 402
              :   |||:::  :  ||| |   |:|| ||  |   : ||||     | :  :| |
Db        529 RRVRGVVELLLARKISVNAKDEDQWTALHFAAQNGDESSTRLLLEKNASVNEVDFEGRTP 588

Qy        403 ---ALQMGNEN-VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEG 458
                | | || |: ||: |: :                                    ::|
Db        589 MHVACQHGQENIVRILLRRGVDV--------------------------------SLQG 615

Qy        459 RQS-TPLHFAAGYNRVSVVEYLL-QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHG 516
              : :  |||:||    : :|: |  | | |  |:|:   |  ||| |   ||| || :|:
Db        616 KDAWLPLHYAAWQGHLPIVKLLAKQPGVSVNAQTLDGRTPLHLAAQRGHYRVARILIDLC 675

Qy        517 AVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDL 576
              : |||   |||| ||   |     :|||  ||       || | | |
Db        676 SDVNVCSLLAQTPLHVAAETGHTSTARLLLHRGAGKEAMTSDGYTALHL----------- 724

Qy        577 LRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHG 636
                    ||: | || ||  -                                    |::
Db        725 ---------AARNGHLATVK--------------------------------LLVEEK 741

Qy        637 ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCA 696
              ||| |:       || ||:|| :|   |:  :  ::  |:   : || |||    |
Db        742 ADVLARGPLNQTALHLAAAHGHSEVVEELVSADV-IDLFDEQGLSALHLAAQGRHAQTVE 800

Qy        697 LLLAHGAD---PTLKNQEGQTP 715
               || |||     :|| | |  |
Db        801 TLLRHGAHINLQSLKFQGGHGP 822


RESULT 15
V244_FOWPV
ID   V244_FOWPV     STANDARD;      PRT;   668 AA.
AC   Q9J4Z6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV244.
GN   FPV244.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 12 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF198100; AAF44588.1; -.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 10.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       40     69       ANK 1.
FT   REPEAT      144    173       ANK 2.
FT   REPEAT      177    206       ANK 3.
FT   REPEAT      210    239       ANK 4.
FT   REPEAT      272    302       ANK 5.
FT   REPEAT      306    336       ANK 6.
FT   REPEAT      340    370       ANK 7.
FT   REPEAT      374    403       ANK 8.
FT   REPEAT      407.   437       ANK 9.
FT   REPEAT      441    471       ANK 10.
FT   REPEAT      473    502       ANK 11.
FT   REPEAT      571    602       ANK 12.
SQ   SEQUENCE   668 AA;  76200 MW;  F3C5BA0760023791B CRC64;

  Query Match             6.6%;  Score 380.5;  DB 1;  Length 668;
  Best Local Similarity   27.3%;  Pred. No. 1.9e-15;
  Matches  153;  Conservative   95;  Mismatches  192;  Indels  121;  Gaps   26;

Qy        186 CSYGHYEVTELLV-----KHGA--CVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPT 238
              |    : |:   :       ||||  | |  :   || :|:|    ::::   |:    :|
Db         11 CRGNYIEILSAITNYNLHKHGANQCEN--ESIPFTAIHQALQLRQIDIVKELIQ--QNPK 66

Qy        239 LL----NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              |:    :  | :   :  || : :          ::    : |:    |: |: ::| ||
Db         67 LIYVTDHRRNSTLHTICITPNVMD---------IVISLTVDCDIILDIKYASI-ILN-KH 115

Qy        295 PQTHETALHCAAASPYPKRKQICELLLRKGANINE----KTKEFLTPLHVASEKAHND-- 348
                                |  : |  :|::| : ||    |  : :  :  :|:   |
Db        116 -----------------KLGEACIHVLKEGISGNEISYNKINKSIEYMKLIKERIQQDEL 158

Qy        349 -VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ-- 405
               : |:::|    ||| |   :| :|  ||   |: :    ||||||  | |||:|   : |:
Db        159 LIAEMLLKKGIDVNAKDVYCRTPIHYAAERGNTKMVNLLLSYGADVNIITLDDLSVLEYA 218

Qy        406 MGNENVQQLLQEGISLGN-SEADRQLLEAAKAGDVETVKKLC-TVQSVNCRDIEGRQSTP 463
              : ::|:  :      :  | :: |  ||:| :  |:||    |  :  |||  |:   ::||
Db        219 VDSKNIDTIKAIIDNRSNINKNDLSLLKAIRNTDLETSLLLYDSGFSVNSIDV--YKNTP 276

Qy        464 LHFAAGYNRVS-VVEYLLQHGADVHAKDKGGLVPLHNACSYGH-YEVAELLVKHGAVVNV 521
              ||:|      :| :|  ||: | ||:||:  |  ||:      |:  |      |:  || ||
Db        277 LHYAVQAPSLSRLVPKLLERGIDVNAKNIKGETPLYLMAKNGYDTENIRTLIMRGADVNA 336

Qy        522 ADLWKFTPLHEAAAKGKYEICKL-LLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 580
              ||    ||||:|:   :|:   : ||: ||
Db        337 ADSLYITPLHQASTLDRYKDTVITLLELGA---------------------------- 366

Qy        581 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 640
                                  ||| ||   :  ||:| ||  ||: :   || :|||:
Db        367 --------------------NVNARDYCDK--TPIHYAAVRNNVVIINTLLDYGADIE 402

Qy        641 A-QDKGGLIPLHNAASYG---HVDVAALLIKYNACVNATDKWAFTPLHEAAQKG-RTQLC 695
              |   | | : ||  | ||    :: |    ||    | ||: :|:  |||| | :|   : ::
Db        403 ALSQKIGTV-LH-FALYGTNPYMSVKT-LIDRGANVNSKNKYLSTPLHYACKKNCKPEVI 459

Qy        696 ALLLAHGADPTLKNQEGQTPL 716
               :|| :|||     |   | ||
Db        460 KMLLDNGADVNAINIRNQYPL 480


RESULT 16
DAK1_HUMAN
ID   DAK1_HUMAN     STANDARD;      PRT;  1432 AA.
AC   P53355; Q9BTL8;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Death-associated protein kinase 1 (EC 2.7.1.37) (DAP kinase 1).
GN   DAPK1 OR DAPK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, INDUCTION, AND MUTAGENESIS OF LYS-42.
RX   MEDLINE=95129831; PubMed=7828849;
RA   Deiss L.P., Feinstein E., Berissi H., Cohen O., Kimchi A.;
RT   "Identification of a novel serine/threonine kinase and a novel 15-kD
RT   protein as potential mediators of the gamma interferon-induced cell
RT   death.";
RL   Genes Dev. 9:15-30(1995).
RN   [2]
RP   REVISIONS TO 164-171.
RA   Feinstein E.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-367 FROM N.A.
RA   Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA   Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA   Phelan M., Farmer A.;
RT   "Cloning of human full-length CDSs in BD Creator(TM) system donor
RT   vector.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=20094983; PubMed=10629061;
RA   Inbal B., Shani G., Cohen O., Kissil J.L., Kimchi A.;
RT   "Death-associated protein kinase-related protein 1, a novel
RT   serine/threonine kinase involved in apoptosis.";
RL   Mol. Cell. Biol. 20:1044-1054(2000).
RN   [5]
RP   FUNCTION, ENZYME REGULATION, MUTAGENESIS OF LYS-42; SER-308 AND
RP   SER-313, AND PHOSPHORYLATION OF SER-308.
RX   PubMed=11579085;
RA   Shohat G., Spivak-Kroizman T., Cohen O., Bialik S., Shani G.,
RA   Berissi H., Eisenstein M., Kimchi A.;
RT   "The pro-apoptotic function of death-associated protein kinase is
RT   controlled by a unique inhibitory autophosphorylation-based
RT   mechanism.";
RL   J. Biol. Chem. 276:47460-47467(2001).
CC   -!- FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase
CC       which acts as a positive regulator of apoptosis.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- COFACTOR: Magnesium.
CC   -!- ENZYME REGULATION: Negatively regulated by autophosphorylation on
CC       Ser-308.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic. Colocalizes with the actin
CC       filament system.
CC   -!- INDUCTION: Up-regulated following treatment with interferon-gamma.
CC   -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
CC       kinase subfamily.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to
CC       frameshifts in positions 466 and 477.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X76104; CAA53712.1; ALT_FRAME.
DR   EMBL; BT006935; AAP35581.1; ALT_TERM.
DR   PDB; 1JKT; 01-APR-02.
DR   Genew; HGNC:2674; DAPK1.
```

```
DR   MIM; 600831; -.
DR   GO; GO:0004684; F:calmodulin-dependent protein kinase I activity; TAS.
DR   GO; GO:0008624; P:induction of apoptosis by extracellular sig. . .; TAS.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF00023; ank; 8.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 7.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW   Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW   3D-structure.
FT   DOMAIN       13    275       PROTEIN KINASE.
FT   DOMAIN      267    334       CALMODULIN-BINDING.
FT   REPEAT      378    407       ANK 1.
FT   REPEAT      411    440       ANK 2.
FT   REPEAT      444    475       ANK 3.
FT   REPEAT      479    508       ANK 4.
FT   REPEAT      512    541       ANK 5.
FT   REPEAT      545    574       ANK 6.
FT   REPEAT      578    607       ANK 7.
FT   REPEAT      611    640       ANK 8.
FT   REPEAT      877    906       ANK 9.
FT   REPEAT     1164   1198       ANK 10.
FT   DOMAIN     1314   1398       DEATH.
FT   NP_BIND      19     27       ATP (BY SIMILARITY).
FT   BINDING      42     42       ATP.
FT   ACT_SITE    139    139       BY SIMILARITY.
FT   MOD_RES     308    308       PHOSPHORYLATION (AUTO-).
FT   MUTAGEN      42     42       K->A: LOSS OF ACTIVITY, APOPTOTIC
FT                                FUNCTION AND OF AUTOPHOSPHORYLATION.
FT   MUTAGEN     308    308       S->A: ELEVATED CA(2+)-CALMODULIN BINDING
FT                                AND CA(2+)-CALMODULIN-INDEPENDENT KINASE
FT                                ACTIVITY. INCREASES APOPTOTIC ACTIVITY.
FT   MUTAGEN     308    308       S->D: REDUCED CA(2+)-CALMODULIN BINDING
FT                                AND CA(2+)-CALMODULIN-INDEPENDENT KINASE
FT                                ACTIVITY. DECREASES APOPTOTIC ACTIVITY.
FT   MUTAGEN     313    313       S->A: MINIMAL EFFECT ON ACTIVITY.
FT   CONFLICT    364    366       LGS -> SKR (IN REF. 3).
SQ   SEQUENCE   1432 AA;  160220 MW;  9B88B66C21CAC16C CRC64;

  Query Match             6.4%;  Score 370;  DB 1;  Length 1432;
  Best Local Similarity   22.3%;  Pred. No. 2.6e-14;
  Matches  225;  Conservative  138;  Mismatches  366;  Indels  278;  Gaps   40;

Qy         58 NYTPLHEAAIKGKIDVCIVL-LQHGAE--PTIRNTDGRTALDLADPSAKAVLTGEYKKDE 114
              |   |||    : | || ::| |  | |     :   :  |  : |     | :| | ||
Db         75 NVITLHE-VYENKTDVILILELVAGGELFDFLAEKESLTE-EEATEFLKQILNGVYYLHS 132

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |  :     | :| | |           |      |       |:||:   |  | |   :
Db        133 LQIAHFDLKPENIMLL-----------DRNVPKP--------RIKIIDFGLAHKIDFGNE 173

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              |           :| |              | ::  : ||    |     : | : :| |
Db        174 FKN---------IFGTPEFV-----------APEIVNYEPLGLEADMWSIGVITYILLSG 213

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH--SLLQAAREADVTRI-----KKHLSL 287
              | | | :   ::  :::        : |||:    |   | :  : |:      || :::
Db        214 ASPFLGDTKQETLANVS-------AVNYEFEDEYFSNTSALAKDFIRRLLVKDPKKRMTI 266

Qy        288 EMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN-EKTKEFLTPLHVASEKAH 346
              :          ::  |     |: | |   : | || : :| || |:|      |: |
Db        267 Q----------DSLQH-----PWIKPKDTQQALSRKASAVNMEKFKKF------AARKKW 305

Qy        347 NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR-------------LLLSYGCD 392
              | :       ::    | :: | |:    ::   |                ::    |
Db        306 KQSVRL-------ISLCQRLSRSFLSRS----NMSVARSDDTLDEEDSFVMKAIIHAIND 354

Qy        393 PNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQS-V 451
              |:  ||      : | :| |           |      || ||  |::: :: |    | :
Db        355 DNVPGLQHLLG-SLSNYDVNQ--------PNKHGTPPLLIAAGCGNIQILQLLIKRGSRI 405

Qy        452 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVA-- 509
              : :|  | |  :::|| :  |  :::| ::   :  ||| | : || |  ||| :||
Db        406 DVQDKGG--SNAVYWAARHGHVDTLKFLSENKCPLDVKDKSGEMALHVAARYGHADVAQV 463

Qy        510 ELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDG 569
              ::|   |:  |: |  : |||| ||  | | : | | : |  :    |||:| ||
Db        464 QVLCSFGSNPNIQDKEEETPLHCAAWHGYYSVAKALCEAGCNVNIKNREGETP------- 516

Qy        570 DTDIQDLLRGDAALLDAAKKG------CLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGY 623
                    || |: :|      |||          ::|  |  |      ||||
Db        517 -------------LLTASARGYHDIVECLA-----EHGADLNACDKDGH--IALHLAVRR 556

Qy        624 NNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPL 683
              :|| : ||  |  |: ||: |  |||  |    |:: :   | : |  :: ::|:  |||
Db        557 CQMEVIKTLLSQGCFVDYQDRHGNTPLHVACKDGNMPIVVALCEANCNLDISNKYGRTPL 616

Qy        684 HEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADD---VSALL------------ 728
              | ||  |   :   |   ||        :|:|  ||  ::    |: ||
Db        617 HLAANNGILDVVRYLCLMGASVEALTTDGKTAEDLARSEQHEHVAGLLARLRKDTHRGLF 676

Qy        729 ------TAAMPP--------------SALPSCYKPQVLNGV---RSPGATADALSSGPSS 765
                    |  : |              : |   |  :|      | |  ::   |  | |
Db        677 IQQLRPTQNLQPRIKLKLFGHSGSGKTTLVESLKCGLLRSFFRRRRPRLSSTNSSRFPPS 736

Qy        766 P--SSLSAASSLDNLSGSFSELS----SVVSSSG-TEG---------------------- 796
              |  |  : :  |::||          :|    |::    | |:|
Db        737 PLASKPTVSVSINNLYPGCENVSVRSRSMMFEPGLTKGMLEVFVAPTHHPHCSADDQSTK 796

Qy        797 ASSLEKKEVPGV-DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN--- 852
              |  ::   : || |||: :|  |       | |      |   :|    :|  || :|
Db        797 AIDIQNAYLNGVGDFSVWEFSGNPVYFCCYDYFAANDPTSIHVVVFSLEEPYEIQLNPVI 856

Qy        853 ---------------AYGHR------------HKLIKGVERLISGQQGL---------- 874
                            |:| :            | | | | |   |: |
Db        857 FWLSFLKSLVPVEEPIAFGGKLKNPLQVVLVATHADIMNVPRPAGGEFGYDKDTSLLKEI 916

Qy        875 -----------NPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV 910
                        |    |:  ||      | | |  ::  :|::| :
Db        917 RNRFGNDLHISNKLFVLDAGASG------SKDMKVLRNHLQEIRSQI 957


RESULT 17
V024_FOWPV
ID   V024_FOWPV     STANDARD;      PRT;   596 AA.
AC   Q9J5H7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV024.
GN   FPV024.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
```

```
CC   -!- SIMILARITY: Contains 13 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44368.1; -.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 12.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT        5     34       ANK 1.
FT   REPEAT       37     66       ANK 2.
FT   REPEAT       68     96       ANK 3.
FT   REPEAT       99    128       ANK 4.
FT   REPEAT      130    158       ANK 5.
FT   REPEAT      162    191       ANK 6.
FT   REPEAT      195    225       ANK 7.
FT   REPEAT      229    258       ANK 8.
FT   REPEAT      262    291       ANK 9.
FT   REPEAT      295    324       ANK 10.
FT   REPEAT      326    355       ANK 11.
FT   REPEAT      359    389       ANK 12.
FT   REPEAT      394    423       ANK 13.
SQ   SEQUENCE   596 AA;  67938 MW;  F813C43228907D07 CRC64;

  Query Match             6.4%;  Score 368;  DB 1;  Length 596;
  Best Local Similarity   25.2%;  Pred. No. 9.1e-15;
  Matches  135;  Conservative   83;  Mismatches  177;  Indels  140;  Gaps   19;

Qy        146 STPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVN 205
              ||||:||   : : ||: || :|||:   :|    ||| | : |: |: :||| ||| :
Db         39 STPLYLAVSNSDIDIVKFLLDNGADI---NKCKSPPLHKAINLGNVEMVKLLVDHGADIE 95

Qy        206 AMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL------NCHNKSAIDLAPTPQLKER 259
              : |   :||: |  |     :   ||  ||||  |         :||  ||:
Db         96 KVYLGN-SPLYLALCKRNTNITKYLLERGADPNTLFINYCDAIYNKIPIDI--------- 145

Qy        260 LAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICEL 319
                ||            : : |  |:::  :|| |  :  |: |   : ||    : :|
Db        146 ----FK----------ILIKYKVSLNIQNSHFKTPIYY--AIKC---TNYP----LIKL 181

Qy        320 LLRKGANINEKTKEFLTPLHVASEKAHN---DVVEVVVKHEAKVNALDNLGQTSLHRAAY 376
              ||   |::     |     :: :   ||    :: :::|:   || |:| :||||
Db        182 LLENNASLT--IPEGYNNHYLITAVKHNCDISILRLLIKYGVPVNEQDDLERTSLHYCVS 239

Qy        377 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 436
              |    :||| |  ||||     | :|      |   :|  :  |   |:|  |
Db        240 AGKHDILKLLLDYDADPNI------TDSCLGTP-----LHYAVSRNDIIATTLLIE--KG 286

Qy        437 GDV----ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 492
              :|    :|:                   | |: | |     ::  || :||:   | :
Db        287 ANVNIHNDTI-----------------DTVLNIAVGNRNKILINLLLMYGANTRLKSRN 328

Qy        493 GLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAK-GKYEICKLLLQHGAD 551
              |:  | |       :   :: ||| ||: :   :|||: |     : :  ||||::|::
Db        329 PLI--HKALETKDINILSEILNHGAEVNIYNREGYTPLYIAIITFMQIKFAKLLLRYGSN 386

Qy        552 PTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQG 611
              |  ||                                                :
Db        387 PNMKN-------------------------------------------------ES 393

Qy        612 RHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLI 666
              :|||| |   | |:  | |: :  ||:: :| |  || :  :|   :| ::|
Db        394 NENTPLHGAILSNRLDSVELLMSYNVDVHSINKLGHTPL-SCINYISDKIATIII 447
```

```
RESULT 18
DAK1_MOUSE
ID   DAK1_MOUSE     STANDARD;      PRT;  1442 AA.
AC   Q80YE7; Q80YE6; Q8R341; Q8VDN6; Q9CSD4; Q9JJP7;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Death-associated protein kinase 1 (EC 2.7.1.37) (DAP kinase 1).
GN   DAPK1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND TISSUE
RP   SPECIFICITY.
RC   TISSUE=Embryo;
RX   MEDLINE=21523887; PubMed=11485996;
RA   Jin Y., Blue E.K., Dixon S., Hou L., Wysolmerski R.B., Gallagher P.J.;
RT   "Identification of a new form of death-associated protein kinase that
RT   promotes cell survival.";
RL   J. Biol. Chem. 276:39667-39678(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RA   Kimchi A.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6; TISSUE=Brain, and Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 1203-1442 FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
```

```
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Calcium/calmodulin-dependent serine/threonine kinase.
CC       Isoform 1 is a negative regulator of apoptosis in direct contrast
CC       to its human homolog. Isoform 2 has no effect on apoptosis and its
CC       function is unknown.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- COFACTOR: Magnesium.
CC   -!- ENZYME REGULATION: Negatively regulated by autophosphorylation on
CC       Ser-308 (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=Beta;
CC         IsoId=Q80YE7-1; Sequence=Displayed;
CC       Name=2; Synonyms=Alpha;
CC         IsoId=Q80YE7-2; Sequence=VSP_050629;
CC   -!- TISSUE SPECIFICITY: High levels in bladder, uterus, vas deferens,
CC       lung, liver and kidney.
CC   -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. DAP
CC       kinase subfamily.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY245540; AAO91934.2; -.
DR   EMBL; AY245541; AAO91935.1; -.
DR   EMBL; X97048; CAA65762.1; -.
DR   EMBL; BC021490; AAH21490.1; ALT_INIT.
DR   EMBL; BC026671; AAH26671.1; -.
DR   EMBL; BC057317; AAH57317.1; -.
DR   EMBL; BC060161; AAH60161.1; -.
DR   EMBL; AK013153; BAB28681.1; -.
DR   MGD; MGI:1916885; Dapk1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   Pfam; PF00023; ank; 8.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   Transferase; Serine/threonine-protein kinase; Calmodulin-binding;
KW   Phosphorylation; ATP-binding; Repeat; ANK repeat; Apoptosis;
KW   Alternative splicing.
FT   DOMAIN       13    275       PROTEIN KINASE.
FT   DOMAIN      267    334       CALMODULIN-BINDING (BY SIMILARITY).
FT   REPEAT      378    407       ANK 1.
FT   REPEAT      411    440       ANK 2.
FT   REPEAT      444    473       ANK 3.
FT   REPEAT      477    506       ANK 4.
FT   REPEAT      510    539       ANK 5.
FT   REPEAT      543    572       ANK 6.
FT   REPEAT      576    605       ANK 7.
FT   REPEAT      609    638       ANK 8.
FT   REPEAT      875    904       ANK 9.
FT   REPEAT     1164   1196       ANK 10.
FT   DOMAIN     1312   1396       DEATH.
FT   NP_BIND      19     27       ATP (BY SIMILARITY).
FT   BINDING      42     42       ATP (BY SIMILARITY).
FT   ACT_SITE    139    139       BY SIMILARITY.
FT   MOD_RES     308    308       PHOSPHORYLATION (AUTO-) (BY
FT                                SIMILARITY).
FT   VARSPLIC   1431   1442       Missing (in isoform 2).
FT                                /FTId=VSP_050629.
FT   CONFLICT    354    354       D -> N (IN REF. 2).
FT   CONFLICT    441    441       K -> Q (IN REF. 3;
FT                                AAH57317/AAH60161).
FT   CONFLICT    461    461       V -> A (IN REF. 2).
FT   CONFLICT    960    960       S -> P (IN REF. 2).
FT   CONFLICT   1000   1000       N -> T (IN REF. 2).
FT   CONFLICT   1038   1042       RWLCT -> PMALH (IN REF. 2).
FT   CONFLICT   1105   1105       A -> V (IN REF. 1).
SQ   SEQUENCE   1442 AA;  161440 MW;  243A14D7C6598F63 CRC64;

  Query Match             6.3%;  Score 364;  DB 1;  Length 1442;
  Best Local Similarity   21.8%;  Pred. No. 6e-14;
  Matches  218;  Conservative  140;  Mismatches  375;  Indels  268;  Gaps   37;

Qy         58 NYTPLHEAAIKGKIDVCIVL-LQHGAE--PTIRNTDGRTALDLADPSAKAVLTGEYKKDE 114
              |   |||    : | || ::| |  | |     :   :  |  : |     | :|:| |
Db         75 NVITLHE-VYENKTDVILILELVAGGELFDFLAEKESLTE-EEATEFLKQILSGVYYLHS 132

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              | :     | :| |                |     |          |:||:   | | |   :
Db        133 LQIAHFDLKPENIMLL-----------DRNVPKP--------RIKIIDFGLAHKIDFGNE 173

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              |             :|  |              | ::  : ||      |   : | : :|  |
Db        174 FKN---------IFGTPEFV-----------APEIVNYEPLGLEADMWSIGVITYILLSG 213

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVN--- 291
              | | |  :   ::  :::         : |:|           | :  |    |: : :
Db        214 ASPFLGDTKQETLANVS-------AVNYDF----------EEEFFRNTSTLAKDFIRRLL 256

Qy        292 FKHPQTHETALHCAAASPYPKRKQICELLLRKGANIN-EKTKEFLTPLHVASEKAHNDVV 350
              | |:   | :  :   |: | |   : | || : :| || |:|       |: |      |
Db        257 VKDPKKRMT-IQDSLQHPWIKPKDTQQALSRKASAVNMEKFKKF------AARKKWKQSV 309

Qy        351 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR--------------LLLSYGCDPNII 396
              :       ::    | :: | |:    ::   |                 ::    | |:
Db        310 RL-------ISLCQRLSRSFLSRS----NMSVARSDDTLDEEDSFVMKAIIHAINDDNVP 358

Qy        397 SLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQS-VNCRD 455
              ||          :  |   ::  |       || ||  |::: :: |     | :: :|
Db        359 GLQHL---------LGSLSSYDVNQPNKHGTPPLLIAAGCGNIQMLQLLIKRGSRIDVQD 409

Qy        456 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 515
              |  |  :::|: :  |  :::| ::   :  ||| |    || |  ||| :| ::||
Db        410 KGG--SNAIYWASRHGHVDTLKFLNENKCPLDVKDKSGETALHVAARYGHADVVQLLCSF 467

Qy        516 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD 575
              |:  :  |  : |||| ||  | | : | | : | :   |||:| ||
Db        468 GSNPDFQDKEEETPLHCAAWHGYYSVAKALCEVGCNVNIKNREGETP------------- 514

Qy        576 LLRGDAALLDAAKKG------CLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVA 629
                    || |: :|       |||          ::|  |  |     ||||     :||
Db        515 -------LLTASARGYHDIVECLA-----EHGADLNASDKDGH--IALHLAVRRCQMEVI 560
```

```
Qy        630 EYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQK 689
              : || ||: |: ||: |  ||| |   |   :   | : :  :: ::|:  |||| ||
Db        561 KTLLGHGSFVDFQDRHGNTPLHVACKDGSAPIVVALCEASCNLDISNKYGRTPLHLAANN 620

Qy        690 GRTQLCALLLAHGADPTLKNQEGQTPLDLVSADD---VSALL----------------- 728
              |   :   |   ||:      :|:|  ||  |:    |: ||
Db        621 GILDVVRYLCLMGANVEALTSDGKTAEDLAKAEQHEHVAGLLARLRKDTHRGLFIQQLRP 680

Qy        729 TAAMPP--------------SALPSCYKPQVLNGV---RSPGATADALSSGPSSP--SSL 769
              |  : |              | |   |  :|      | |  ::   :  | ||  :
Db        681 TQNLQPRIKLKLFGHSGSGKSTLVESLKCGLLRSFFRRRRPRLSSTNSTRFPPSPLAAKP 740

Qy        770 SAASSLDNLSGSFSELS----SVVSSSG-TEG---------------------ASSLEK 802
              : : |::||       :|    |::   | |:|                     |  ::
Db        741 TVSVSINNLYPGCENVSVRSRSMMFEPGLTKGMLEVFVAPSHHLHCSTDDQSTKAIDIQN 800

Qy        803 KEVPGV-DFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN--------- 852
              : || |||: :|  |       | |      |   ::    :|  || :|
Db        801 AYLNGVGDFSVWEFSGNPVYFCCYDYFAANDPTSIHIIVFSLEEPYEIQLNQVIFWLSFL 860

Qy        853 ----------AYGHR------------HKLIKGVERLISGQQGL---------------- 874
                       |:| :            |  |  :  |   |: |
Db        861 KSLVPVEEPIAFGGKLKNPLRVVLVATHADIMNIPRPAGGEFGYDKDTSLLKEIRNRFGN 920

Qy        875 -----NPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV 910
                  |   |:   ||      | | |  ::  :|::| :
Db        921 DLHVSNKLFVLDAGASG------SKDIKVLRNHLQEIRSQI 955
```

RESULT 19
V155_FOWPV
ID V155_FOWPV STANDARD; PRT; 542 AA.
AC Q9J5A7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ankyrin-repeat protein FPV115.
GN FPV115.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- SIMILARITY: Contains 8 ANK repeats.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF198100; AAF44459.1; -.
DR HSSP; P80144; 2MYO.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 8.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 8.
DR PROSITE; PS50088; ANK_REPEAT; 6.
DR PROSITE; PS50297; ANK_REP_REGION; 2.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 33 62 ANK 1.
FT REPEAT 157 186 ANK 2.
FT REPEAT 218 247 ANK 3.
FT REPEAT 251 281 ANK 4.
FT REPEAT 285 314 ANK 5.
FT REPEAT 316 345 ANK 6.
FT REPEAT 347 375 ANK 7.
FT REPEAT 378 407 ANK 8.
SQ SEQUENCE 542 AA; 62056 MW; CBB71D6E04A06629 CRC64;

```
  Query Match              6.1%;  Score 351;  DB 1;  Length 542;
  Best Local Similarity   25.9%;  Pred. No. 8.4e-14;
  Matches  137;  Conservative   85;  Mismatches  171;  Indels  136;  Gaps   24;

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQF--------TPLHEAASKNRVEVCSLLLSY 233
              |||    | |:             : |:::::       |||  |   | ::    ||
Db          6 LHNIMYTGTYD---------EIMRAINIYEHRNYSSFRNLPLHYAIYSRRKDIVETLLKS 56

Qy        234 GADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFK 293
              | ||  ::  : : : |   |        ::    ::|     |     |  :||     | |
Db         57 GYDPNSVDITDNNCLQLLSMP-------FDI---TMLPVDEEVQDYAISFYLS---KNMK 103

Qy        294 H-----PQTHETALHCAAASPYP------------KRKQIC--ELLLRKGANINEKTKEF 334
              |     | | |          : ||           |  ::|  :|||| ||  | : |:
Db        104 HTSMLIPITKE----ALRGNRYPSEPYFSSMCRKFKDNELCIMDLLLRYGALPNSR-KDG 158

Qy        335 LTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRL---LLSYGC 391
              | ||: |:   : ::||::: : || |       :    |    :::| ::    | :|
Db        159 LLPLYHAAAAGNTEMVELLLSYGAKTNLHTRYEDSIFMCAIKSNNVKTAKIISDLYNYKN 218

Qy        392 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 451
              | | |       :|: | ::   |  | :|                          :
Db        219 DINNI----LKTIQLYNADMLLFL---IEIG-------------------------LDI 245

Qy        452 NCRDIEGRQSTPLHFAA-GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAE 510
              | :| :|:  | ||:|    | :  |: ::::|||:: ||: || |||:|| ||  :::: 
Db        246 NTKDKKGK--TALHYACNSINCIETVKEIMKYGADINVKDREGLTPLHSACKYGDLKLSK 303

Qy        511 LLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNT----PLDL- 565
              ||:::|| | |       | |: |   |  |: | |::       || :  |     | |
Db        304 LLIEYGADVKVKT--TSTVLNLAVESGNVELVKFLIE-------KNPEFITSDYLSLSLA 354

Qy        566 VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNN 625
              ::  | :|  ||  ||  :|      |:                      ||||||     |
Db        355 IRCKDINIV-LLLLDAG-MDVNSSKCI--------------------STPLHLGVILGN 391

Qy        626 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAAL------LIKY 668
              : : || |||::|| || |  ||  |    ::| | |       :|||
Db        392 SNIVKLLLDHGANINAIDKYGETPLEAANKRINIDYAELYKSNRFIIKY 440
```

RESULT 20
ASB2_HUMAN
ID ASB2_HUMAN STANDARD; PRT; 587 AA.
AC Q96Q27; Q9NSU5; Q9Y567;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ankyrin repeat and SOCS box containing protein 2 (ASB-2).
GN ASB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21450655; PubMed=11566180;
RA Kohroki J., Fujita S., Itoh N., Yamada Y., Imai H., Yumoto N.,
RA Nakanishi T., Tanaka K.;
RT "ATRA-regulated Asb-2 gene induced in differentiation of HL-60
RT leukemia cells.";
RL FEBS Lett. 505:223-228(2001).
RN [2]
RP SEQUENCE FROM N.A., AND INDUCTION.
RX MEDLINE=21634875; PubMed=11682484;
RA Guibal F.C., Moog-Lutz C., Smolewski P., Di Gioia Y.,
RA Darzynkiewicz Z., Lutz P.G., Cayre Y.E.;
RT "ASB-2 inhibits growth and promotes commitment in myeloid leukemia
RT cells.";

```
RL   J. Biol. Chem. 277:218-224(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 31-587 FROM N.A.
RC   TISSUE=Uterus;
RA   Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA   Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 42-587 FROM N.A.
RX   MEDLINE=20564172; PubMed=11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2000).
CC   -!- FUNCTION: Not known. Its expression induces growth inhibition and
CC       chromatin condensation.
CC   -!- INDUCTION: By all-trans retinoic acid (ATRA).
CC   -!- SIMILARITY: Contains 12 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB056723; BAB64532.1; -.
DR   EMBL; AJ251238; CAC17765.1; -.
DR   EMBL; BC032354; AAH32354.1; -.
DR   EMBL; AL137735; CAB70899.1; -.
DR   EMBL; AF159164; AAD45345.1; -.
DR   PIR; T46507; T46507.
DR   Genew; HGNC:16012; ASB2.
DR   MIM; 605759; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 11.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 9.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       56     85       ANK 1.
FT   REPEAT       89    119       ANK 2.
FT   REPEAT      123    152       ANK 3.
FT   REPEAT      156    185       ANK 4.
FT   REPEAT      189    218       ANK 5.
FT   REPEAT      222    251       ANK 6.
FT   REPEAT      255    284       ANK 7.
FT   REPEAT      288    317       ANK 8.
FT   REPEAT      320    349       ANK 9.
FT   REPEAT      362    391       ANK 10.
FT   REPEAT      392    421       ANK 11.
FT   REPEAT      428    456       ANK 12.
FT   DOMAIN      533    587       SOCS BOX.
FT   CONFLICT    224    224       MISSING (IN REF. 5).
FT   CONFLICT    460    460       MISSING (IN REF. 5).
SQ   SEQUENCE   587 AA;  65083 MW;  866A2B6010C4C1F9 CRC64;

  Query Match             6.1%;  Score 350;  DB 1;  Length 587;
  Best Local Similarity   24.3%;  Pred. No. 1.1e-13;
  Matches  133;  Conservative   72;  Mismatches  190;  Indels  152;  Gaps   15;

Qy         28 PLHNACSFGHAEVVNLLLRHG---ADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEP 84
              ||  |    |  | :  ::: |   |:|| :: |   ||||||  |::  |: :||
Db         60 PLIKAIKDGDEEALKTMIKEGKNLAEPN-KEGW--LPLHEAAYYGQVG-CLKVLQRAYPG 115

Qy         85 TIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGR 144
              ||   | ||   | : :|  : |     | ||   ::| |  :             |:
Db        116 TI---DQRT---LQEETAVYLATCRGHLDCLLSLLQAGAEPDI------------SNKS 156

Qy        145 KSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACV 204
              : |||: |      : |::|:|| || : :       || : |    || ::||  || |
Db        157 RETPLYKACERKNAEAVKILVQHNADTNHRCNRGWTALHESVSRNDLEVMQILVSGGAKV 216

Qy        205 NAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEF 264
              : : :  |||  ||   ::|     | ||||       | ||
Db        217 ESKNAYGITPLFVAAQSGQLEALRFLAKYGADINTQASDNASA---------------- 259

Qy        265 KGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKG 324
                | :| :                                        : ::: | || :|
Db        260 ----LYEACKN-----------------------------------EHEEVVEFLLSQG 279

Qy        325 ANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR 384
              |: |:  |: | |||:||:| :  :|::::   ::   :   | : || ||   | :
Db        280 ADANKTNKDGLLPLHIASKKGNYRIVQMLLPVTSRTR-IRRSGVSPLHLAAERNHDEVLE 338

Qy        385 LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKK 444
              |||   | |
Db        339 ALLSARFDVN---------------------------------------------------- 348

Qy        445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 504
                :       |  | |:|: |:||   | |   | |||||||    ::  : ||  |  :|
Db        349 -TPLAPERARLYEDRRSSALYFAVVNNNVYATELLLQHGAD---PNRDVISPLLVAIRHG 404

Qy        505 HYEVAELLVKHGAVVNVADLWKFTPLHEAA-------AKGKYEICKLLLQHGAD--PTKK 555
                :||: ||| ::      :  | |       |      : | |:  | |  |
Db        405 CLRTMQLLLDHGANIDA-----YIATHPTAFPATIMFAMKCLSLLKFLMDLGCDGEPCFS 459

Qy        556 NRDGNTP 562
                || |
Db        460 CLYGNGP 466


RESULT 21
ANS1_MOUSE
ID   ANS1_MOUSE     STANDARD;      PRT;  1150 AA.
AC   P59672;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ankyrin repeat and SAM domain containing protein 1.
GN   ANKS1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Limb;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 6 ANK repeats.
CC   -!- SIMILARITY: Contains 1 PID domain.
CC   -!- SIMILARITY: Contains 2 SAM domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; BC050847; AAH50847.1; -.
DR   MGD; MGI:2446180; Anks1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR006020; PTB_PID.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 7.
DR   Pfam; PF00640; PID; 1.
DR   Pfam; PF00536; SAM; 2.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00462; PTB; 1.
DR   SMART; SM00454; SAM; 2.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS01179; PID; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 2.
KW   ANK repeat; Repeat.
FT   DOMAIN       31     51       GLY-RICH.
FT   REPEAT       75    104       ANK 1.
FT   REPEAT      108    137       ANK 2.
FT   REPEAT      165    194       ANK 3.
FT   REPEAT      198    227       ANK 4.
FT   REPEAT      231    260       ANK 5.
FT   REPEAT      263    292       ANK 6.
FT   DOMAIN      712    778       SAM 1.
FT   DOMAIN      786    853       SAM 2.
FT   DOMAIN      952   1107       PID.
SQ   SEQUENCE   1150 AA;  125241 MW;  551AF44D1DB02D1B CRC64;

  Query Match             5.9%;  Score 342;  DB 1;  Length 1150;
  Best Local Similarity   19.3%;  Pred. No. 9.1e-13;
  Matches  189;  Conservative  120;  Mismatches  264;  Indels  406;  Gaps   28;

Qy        111 KKDELLESARSGNEEKMMALLTPLNVNCHASD----------------GRKSTPLHLAAG 154
              |: ||||:||:|:   :  ||:   ::                    |  | ||
Db          3 KEQELLEAARTGHLPAVEKLLSGKRLSSGFGGGGGGSGSGGGSGGGGLGSSSHPLSSL-- 60

Qy        155 YNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTP 214
                    | :  | :|:  |      |||:|   || :| |:|::: |  |  |      |
Db         61 --------LSMWRGPNVNCVDSTGYTPLHHAALNGHRDVVEVLLRNDALTNVADSKGCYP 112

Qy        215 LHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAR 274
              || || |   ::  ||:  |   | :|  |
Db        113 LHLAAWKGDAQIVRLLIQQGPSHTRVNEQN---------------------------- 142

Qy        275 EADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEF 334
                ::  :||:                                ||                ||  |  :
Db        143 ALEIRELKKY-------------------GPFDPY----------------INAKNNDN 166

Qy        335 LTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN 394
              | || |::  | :||: :::         :|  :| |  ||  | |:  :|||  |  ||
Db        167 ETALHCAAQYGHTEVVKALLEELTDPTMRNNKFETPLDLAALYGRLEVVKLLL--GAHPN 224

Qy        395 IISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCR 454
              ::|                                                    |:
Db        225 LLS-------------------------------------------------CST------ 230

Qy        455 DIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVK 514
                |: |||| ||     :||: ||  | | : : |
Db        231 ----RKHTPLHLAARNGHKAVVQVLLDAGMDSNYQTEMG---------------------- 265

Qy        515 HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTD-- 572
                      : ||||  || :: ::||  | |   |:  | | || |:|  :
Db        266 -------------SALHEAALFGKTDVVQILLAAGIDVNIKDNRGLTALDTVRDLPSQKS 312

Qy        573 ------IQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 626
                   |:| : |  :: :   :    |::  ||: | : |              :: |
Db        313 QQIAALIEDHMTGKRSVKEVDRTS-TAQLPLLSNTDAI------------APMSQGSMEK 359

Qy        627 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAA---------------------- 663
              | | :|    | :| ::|    |:|| |   :|  |
Db        360 TVTELILHF--DTHADEEGPYEALYNAVSCHSLDSTASGRSSDRDSMNKEAEATGTRAAG 417

Qy        664 ------------------------------------------------------------LLI 666
                                                                           ||
Db        418 VRPRERPPPPAKPPPDEEEEERVDKKYFPLAASEGLAVRPRIQSSAPQEEEEHPYELLLT 477

Qy        667 KYNACVNATDKWAFTPLHEAAQKGRT-------------QLCALLLAHGADPTLKNQEGQ 713
                  :  ||    |  |  :  ||:              |   ||  ||  ||: |    : ||
Db        478 AETKKLGTTD--GRTEDHRQSGSGRSQDSVEGQDGQVPEQFSGLL--HGSSPVC--EVGQ 531

Qy        714 TPLDLVSADDVS-------------------------------------------ALLTA- 730
              |  |::|   |                                           | |||
Db        532 DPFQLLTAPSQSHPESSQQDACHEASMQLEEPGVQGTEPPQPGVPDQSKRVGLPAGLTAL 591

Qy        731 ------------------------AMPPSALPSCYKPQV-LNGVRS-PGATADALSSGP 763
                                       : | |  |   ||:  |   ||   : :| |:  |
Db        592 ASRTYLDALTHTVPLRPAGAEEEDQSGPRSRAPPTSKPKAELKLSRSLSKSDSDLLTCSP 651

Qy        764 SSPSSL-SAASSLDNLS-------------GSFSELSSVVSSSGTEGASSLEKKEVPG-- 807
              :  ::: | : || | |               : |:  ::|| |      | |:::: |
Db        652 TEDATMGSRSESLSNCSIGKKRLEKSPSFASEWDEIEKIMSSIGEGIDFSQEQQKISGSR 711

Qy        808 -VDFSITQFVRNLGLEH-----LMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLI 861
               :: |: ::: ::||:      |:: |:  :     ::|   ::|:||||:   || ||:
Db        712 TLEQSVGEWLESIGLQQYESKLLLNGFDDVRFLGSNVME--EQDLREIGISDPQHRRKLL 769

Qy        862 KGVERL----ISGQQGLNP 876
              :    |      |  |::|
Db        770 QAARSLPKVKALGYDGVSP 788


RESULT 22
ANR5_MOUSE
ID   ANR5_MOUSE     STANDARD;      PRT;   775 AA.
AC   Q9D2J7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat domain protein 5.
```

```
GN   ANKRD5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK019547; BAB31791.1; -.
DR   HSSP; P80144; 2MYO.
DR   MGD; MGI:2441685; Ankrd5.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT      184    213       ANK 2.
FT   REPEAT      217    246       ANK 3.
FT   REPEAT      250    279       ANK 4.
FT   REPEAT      524    553       ANK 5.
FT   REPEAT      557    586       ANK 6.
FT   REPEAT      590    619       ANK 7.
FT   REPEAT      623    652       ANK 8.
SQ   SEQUENCE   775 AA;  86903 MW;  51C11E1BC5151F00 CRC64;

  Query Match             5.8%;  Score 334.5;  DB 1;  Length 775;
  Best Local Similarity   24.7%;  Pred. No. 1.4e-12;
  Matches  172;  Conservative   86;  Mismatches  270;  Indels  169;  Gaps   26;

Qy        115 LLESARSGNE---EKMMALLTPLNVN-CHASDGRKSTPLHLAAGYNRVKIVQLLLQHGAD 170
              :|:  |: ::   ||:  |  |  :|     ||  :  ||||:  |   :|  ||: ||
Db         16 VLQCVRNKDKKQIEKLTRLGYPELINFTEPIDGLSA--LHLASISNDTDMVSFLLKLGAH  73

Qy        171 VHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD-------LWQFTPLHEAASKNR 223
                :|     |  |   ||    |:| |  | :  :|       :  |        |
Db         74 PDVQDHMGCTPTMRAAELGHELSMEILAKAKADMTIVDNEGKGVLFYCILP------TKR 127

Qy        224 VEVCSLL-LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIK 282
                ||:  | :|||                    : || |    |:|  ||
Db        128 HYRCSLIALEHGAD------------------VNNITYEGK-PVFLRACEEA------ 160

Qy        283 KHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINE-KTKEFLTPLHVA 341
                          |:                | :|   | |||| |   |    | | :
Db        161 --------------HDV------------KDMCLTFLEKGANPNAINTSTGRTALMES 192

Qy        342 SEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGF 401
              | :  ::|  :::   :||| ||    : | ||  |     :|| :|  |  :| : |
Db        193 SREGVLEIVRGILERGGEVNAYDNDRHHAAHFAAKGGFFDILKLLFAYNGDMGLIGMDGN 252

Qy        402 TALQ---MGN--ENVQQLLQEGISL--GNSE---------------------ADRQLLE 432
              | |    ||   :  : : | |  |   | |                     |:|  :
Db        253 TPLHFAAMGGFADCCKYIAQRGCDLKWKNLEHKTPRVVAKDGGFKAASKEIRRAERTAAK 312

Qy        433 AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 492
              || |     |    : ::   |     | |  |  :          :  |  | :||
Db        313 LAKTG----AKNPNPLWALRLHDWSIEHETSLRNAFKF---------VDRGDGVVSKD-D 358

Qy        493 GLVPLHNACSYGHYE----VAELLVK-HGAVVNVADLWK-------------FTP----- 529
              :| |     |  |    ||::  |  |  ||: : :|              | |
Db        359 FVVALEERQEYATSEQLLSVAQMHEKSRGGGVNINEFFKGTKYLSKSYVLGSFGPKKKKR 418

Qy        530 -LHEAAAKGKY----EICKL------------------LLQHGADPTKKNRDGNTPLDL 565
               | :  |||:     || :                  |: :|  : |||      :
Db        419 GLGKKPRKGKFVLPLPICTIPENAFPRRPDGGPPYYMIETYQNVSDSHRFNRDHPPEHPI 478

Qy        566 VKDGDTDIQDLLRGDAALLDAAKKGCLARVKK-LSSPDNVNCRDTQGRHSTPLHLAAGYN 624
               | :  | |  |  | :    | | || :|| : :   |: :|    : ||| :|
Db        479 QDDSEWYIDDPSRVFANISFITKAGDLASLKKAIETGIPVDMKD--NTYKTPLMIACASG 536

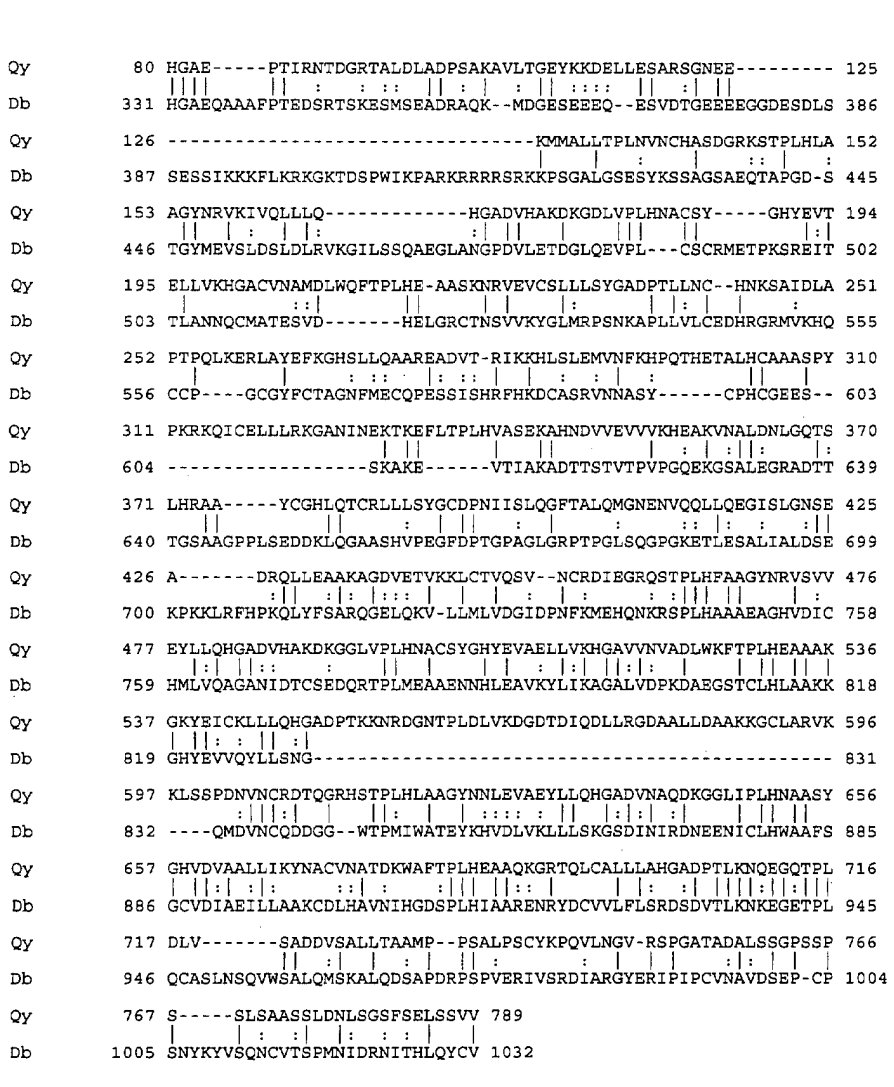
Qy        625 NLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLH 684
              |::| ::|:: ||:||| |     ||| |   | :  ||:|  | ::||     |||
Db        537 NIDVVKFLIEKGANVNATDNFLWTPLHFACHAGQQDIVELLVKAGASIDATSINNSTPLS 596

Qy        685 EAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 721
               | :  |    ||  ||   ::|::|   :|:  |
Db        597 RAIESCRLDTVKYLLDMGAKFQIENRKGHAAMDIAKA 633


RESULT 23
HMT1_HUMAN
ID   HMT1_HUMAN     STANDARD;      PRT;  1267 AA.
AC   Q9H9B1; Q8TCN7; Q96F53; Q96JF1; Q96KH4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Histone-lysine N-methyltransferase, H3 lysine-9 specific 5
DE   (EC 2.1.1.43) (Histone H3-K9 methyltransferase 5) (H3-K9-HMTase 5)
DE   (Euchromatic histone methyltransferase 1) (Eu-HMTase1) (G9a-like
DE   protein 1) (GLP1).
GN   EUHMTASE1 OR KIAA1876.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), ENZYMATIC ACTIVITY, AND IDENTIFICATION
RP   IN COMPLEX WITH E2F6; TFDP1; MAX; MGA; BAT8; CBX3; RING1; RNF2; MBLR;
RP   L3MBTL2 AND YAF2.
RC   TISSUE=Cervical carcinoma;
RX   MEDLINE=21999559; PubMed=12004135;
RA   Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT   "A complex with chromatin modifiers that occupies E2F- and
RT   Myc-responsive genes in G0 cells.";
RL   Science 296:1132-1136(2002).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Teratocarcinoma;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
```

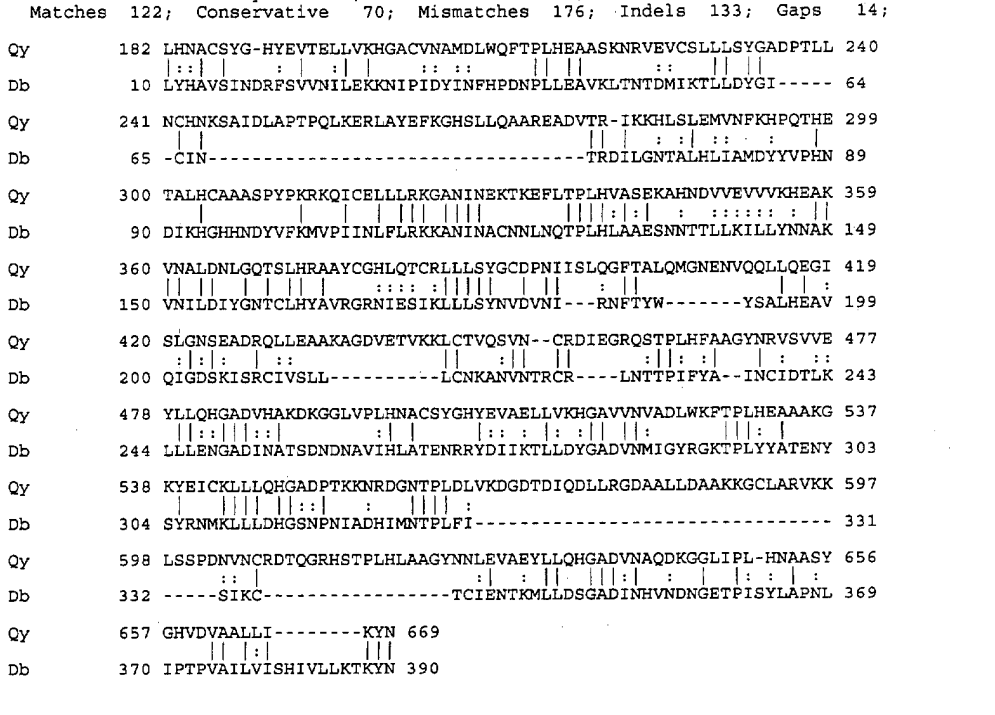
```
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 351-1267 FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RA   Tsuritani K., Ukai Y., Yajima Y., Amemiya C., Yoshimoto M.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 351-1267 FROM N.A. (ISOFORM 3).
RX   MEDLINE=22158633; PubMed=12168954;
RA   Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT   "Construction of expression-ready cDNA clones for KIAA genes: manual
RT   curation of 330 KIAA cDNA clones.";
RL   DNA Res. 9:99-106(2002).
RN   [5]
RP   SEQUENCE OF 1150-1267 FROM N.A. (ISOFORM 3), AND TISSUE SPECIFICITY.
RC   TISSUE=Brain;
RX   MEDLINE=21245130; PubMed=11347906;
RA   Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XX.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 8:85-95(2001).
RN   [6]
RP   SEQUENCE OF 552-1267 FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SEQUENCE OF 1056-1267 FROM N.A. (ISOFORM 1).
RC   TISSUE=Lymph node;
RA   Ansorge W., Wirkner U., Mewes H.-W., Weil B., Wiemann S.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Histone methyltransferase. Methylates Lys-9 of histone
CC       H3 (in vitro). H3 Lys-9 methylation represents a specific tag for
CC       epigenetic transcriptional repression by recruiting HP1 proteins
CC       to methylated histones. Probably targeted to histone H3 by
CC       different DNA-binding proteins like E2F6, MGA, MAX and/or DP1.
CC       During G0 phase, it probably contributes to silencing of MYC- and
CC       E2F-responsive genes, suggesting a role in G0/G1 transition in
CC       cell cycle.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC       S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC   -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC       E2F6, MGA, MAX, TFDP1, CBX3, BAT8, EUHMTASE1, RING1, RNF2, MBLR,
CC       L3MBTL2 and YAF2.
CC   -!- SUBCELLULAR LOCATION: Nuclear; associates with euchromatic
CC       regions.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Experimental confirmation may be lacking for some
CC         isoforms;
CC       Name=1;
CC         IsoId=Q9H9B1-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9H9B1-2; Sequence=VSP_002222, VSP_002223;
CC       Name=3;
CC         IsoId=Q9H9B1-3; Sequence=VSP_002224, VSP_002225;
CC   -!- TISSUE SPECIFICITY: Widely expressed.
CC   -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC       family.
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   -!- SIMILARITY: Contains 1 pre-SET domain.
CC   -!- SIMILARITY: Contains 1 SET domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY083210; AAM09024.1; -.
DR   EMBL; AK022941; BAB14321.1; -.
DR   EMBL; AB028932; BAB56104.1; -.
DR   EMBL; AB058779; BAB47505.2; -.
DR   EMBL; BC011608; AAH11608.2; -.
DR   EMBL; AL713772; CAD28534.1; ALT_INIT.
DR   MIM; 607001; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR007728; Pre-SET.
DR   InterPro; IPR001214; SET.
DR   InterPro; IPR003606; Zn2-binding.
DR   Pfam; PF00023; ank; 7.
DR   Pfam; PF05033; Pre-SET; 1.
DR   Pfam; PF00856; SET; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 7.
DR   SMART; SM00468; PreSET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50867; PRE_SET; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW   ANK repeat; Repeat; Alternative splicing.
FT   REPEAT      706    735       ANK 1.
FT   REPEAT      741    770       ANK 2.
FT   REPEAT      774    803       ANK 3.
FT   REPEAT      807    837       ANK 4.
FT   REPEAT      841    870       ANK 5.
FT   REPEAT      874    903       ANK 6.
FT   REPEAT      907    936       ANK 7.
FT   REPEAT      940    973       ANK 8.
FT   DOMAIN     1029   1092       PRE-SET.
FT   DOMAIN     1094   1216       SET.
FT   DOMAIN      375    378       POLY-GLU.
FT   DOMAIN      411    418       POLY-ARG.
FT   DOMAIN     1261   1264       POLY-ALA.
FT   VARSPLIC    805    825       DAEGSTCLHLAAKKGHYEVVQ -> IQKTSKVYTESQETQR
FT                                SQTIL (in isoform 2).
FT                                /FTId=VSP_002222.
FT   VARSPLIC    826   1267       Missing (in isoform 2).
FT                                /FTId=VSP_002223.
FT   VARSPLIC   1150   1153       DGEV -> ISSA (in isoform 3).
FT                                /FTId=VSP_002224.
FT   VARSPLIC   1154   1267       Missing (in isoform 3).
FT                                /FTId=VSP_002225.
FT   CONFLICT    530    530       G -> E (IN REF. 3 AND 4).
SQ   SEQUENCE   1267 AA;  138181 MW;  78FD418FD404F475 CRC64;

  Query Match             5.7%;  Score 330;  DB 1;  Length 1267;
  Best Local Similarity   22.6%;  Pred. No. 5.6e-12;
  Matches  183;  Conservative  102;  Mismatches  319;  Indels  204;  Gaps   27;
```



RESULT 24
V228_FOWPV
ID V228_FOWPV STANDARD; PRT; 525 AA.
AC Q9J507;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative ankyrin-repeat protein FPV228.
GN FPV228.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- SIMILARITY: Contains 9 ANK repeats.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF198100; AAF44572.1; -.
DR HSSP; P25963; 1IKN.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 9.
DR SMART; SM00248; ANK; 9.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT REPEAT 39 71 ANK 1.
FT REPEAT 72 122 ANK 2.
FT REPEAT 123 152 ANK 3.
FT REPEAT 156 185 ANK 4.
FT REPEAT 190 226 ANK 5.
FT REPEAT 227 254 ANK 6.
FT REPEAT 258 287 ANK 7.
FT REPEAT 291 320 ANK 8.
FT REPEAT 324 353 ANK 9.
SQ SEQUENCE 525 AA; 60155 MW; E8AF61D9B33BCFB0 CRC64;

Query Match 5.7%; Score 327.5; DB 1; Length 525;
Best Local Similarity 24.4%; Pred. No. 2.1e-12;
Matches 122; Conservative 70; Mismatches 176; Indels 133; Gaps 14;



```
RESULT 25
ANR5_HUMAN
ID   ANR5_HUMAN     STANDARD;      PRT;   776 AA.
AC   Q9NU02; Q9H6Y9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat domain protein 5.
GN   ANKRD5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
RN   [2]
RP   SEQUENCE OF 329-776 FROM N.A.
RC   TISSUE=Colon;
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA   Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL109754; CAB87618.2; -.
DR   EMBL; AK025322; BAB15111.1; ALT_INIT.
DR   HSSP; P80144; 2MYO.
DR   Genew; HGNC:15803; ANKRD5.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002048; EF-hand.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT      184    213       ANK 2.
FT   REPEAT      217    246       ANK 3.
FT   REPEAT      250    279       ANK 4.
FT   REPEAT      524    553       ANK 5.
FT   REPEAT      557    586       ANK 6.
FT   REPEAT      590    619       ANK 7.
FT   REPEAT      623    652       ANK 8.
SQ   SEQUENCE   776 AA;  86664 MW;  2F71F35AC4D337B6 CRC64;

  Query Match             5.6%;  Score 322.5;  DB 1;  Length 776;
  Best Local Similarity   22.5%;  Pred. No. 7.6e-12;
  Matches  170;  Conservative  106;  Mismatches  282;  Indels  197;  Gaps   28;

Qy          6 DVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 65
              |:| :|| ||    :|  |  |    |   ||   : :| :  ||    ||
Db         62 DMVSFLLDLGAHPDVQDRMGCTPTMRAAELGHELSMEILAKAKADMTIVDNEG------- 114

Qy         66 AIKGKIDVCI-----------VLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 114
                 || :  ||           : |:|||:      :|:     |  |  |      ||
Db        115 --KGVLFYCILPTKRHYRCALIALEHGADVNNSTYEGKPIFLRACEDAHDV------KDV 166

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |      |          |  :|  :| ||  |  |   |:    |:||: :|: | :|:|
Db        167 CLTFLEKG--------ANPNAIN--SSTGR--TALMEASREGVVEIVRGILERGGEVNAF 214

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              |  |  |  ::: :||  :   |  : :   ||| ||    : |  :    |
Db        215 DNDRHHAAHFAAKGGFFDILKLLFAYNGDVGLISINGNTPLHYAAMGGFADCCKYIAQRG 274

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              |    | :|      ||:                |:|               ||
Db        275 CDLKWKNLDHK-------TPR--------------AVAKEG-------------GFK- 297

Qy        295 PQTHETALHCAAASPYPKRKQICELLLRKGA-NINE----------KTKEFLTPLHVAS 342
                    ||:   : ::|   | | || | |           : : ||
Db        298 ----------AASKEIRRAERIANKLARPGAKNPNPLWALRLHDWSVEREAFLREAPAVL 347

Qy        343 EK-----AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCG-----------HLQTCRL 385
              ::     : || | |: : :   ::       ||     |           :|    :
Db        348 DRGDGSISKNDFVMVLEERQDYASSEQLAAIAHLHEKTRGGGVNINEFFKGTRYLNKSFV 407

Qy        386 LLSYG----------------------CDPNIISLQGFTALQMGN---------ENV-- 411
              | |||                    |  :|     |  | |           :||
Db        408 LGSYGPKKKEKGMGKKGKKGKFVLPLPIC---VIPEYAFPRRQDGGPPYYMIETYKNVTD 464

Qy        412 ----------QQLLQEGISLGNSEADR---QLLEAAKAGDVETVKKLCTVQS---VNCRD 455
                       :  :|:       ::::    :    ||||: ::||    :|   |: :|
Db        465 SSRFNRDHPPEHPIQDDSVWYIDDSEKVFSNINIITKAGDLASLKK--AFESGIPVDMKD 522

Qy        456 IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKH 515
                ||| |     : ||::||: ||:|:| |     ||| ||  |  :: ||||:
Db        523 --NYYKTPLMTACASGNIDVVKFLLEKGANVNATDNFLWTPLHFACHAGQQDIVELLVES 580

Qy        516 GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD 575
              ||::: | :   |||: |    : :  | ||  ||    :|| |:: :|: |
Db        581 GALIDAASINNSTPLNRAIESCRLDTVKYLLDIGAKFQLENRKGHSAMDVAK-------- 632

Qy        576 LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ- 634
                  | ::|  |:    ::  |  |       : : :  ||  |       ::  | ||
Db        633 -AYADYRIIDLIKE----KLDNLPKP----AENQKLKGKTPPILKTEGPEIKKEEELLSS 683

Qy        635 -HGADVNAQ----DKGGLIPLHNAASYGH---VDV 661
               :|    ::      || :: |::  : |:   ||:
Db        684 IYGVPTTSEGKKVQKGNVVHLNSLITSGYTKKVDI 718


RESULT 26
ASB3_MOUSE
ID   ASB3_MOUSE     STANDARD;      PRT;   525 AA.
AC   Q9WV72;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 3 (ASB-3).
GN   ASB3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=20564172; PubMed=11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2000).
CC   -!- TISSUE SPECIFICITY: Widely expressed; highest expression in testis
CC       and spleen.
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF155354; AAD38810.1; -.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:1929749; Asb3.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT        9     38       ANK 1.
FT   REPEAT       42     71       ANK 2.
FT   REPEAT       78    107       ANK 3.
FT   REPEAT      111    140       ANK 4.
FT   REPEAT      145    174       ANK 5.
FT   REPEAT      178    207       ANK 6.
FT   REPEAT      211    240       ANK 7.
FT   REPEAT      246    275       ANK 8.
FT   REPEAT      279    308       ANK 9.
FT   REPEAT      315    346       ANK 10.
FT   DOMAIN      442    505       SOCS BOX.
SQ   SEQUENCE   525 AA;  58354 MW;  FE67680AD3958FC1 CRC64;

  Query Match             5.6%;  Score 322;  DB 1;  Length 525;
  Best Local Similarity   29.7%;  Pred. No. 4.5e-12;
  Matches  118;  Conservative   48;  Mismatches  127;  Indels  104;  Gaps   15;

Qy        433 AAKAGDVETVKKLCTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGAD---VHA 488
              ||: |:|: ::||    :||:  |  |    |:| || :| |  :: |:   :    : |
Db         17 AAREGNVKILRKLLKKGRSVDVADNRGWM--PIHEAAYHNAVECLQMLIHTDSSENYIKA 74

Qy        489 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 548
              |   |   || | | ||::: ::|:: ||  |   | : |||  |   |: :: ||||||
Db         75 KTFEGFCALHLAVSQGHWKITQILLEAGADPNETTLEETTPLFLAVESGRIDVLKLLLQH 134

Qy        549 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 608
              ||                                                      |||
Db        135 GA----------------------------------------------------NVN--- 139

Qy        609 TQGRHS----TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAAL 664
                 | ||      || |:   | |  : ||: |||   ||   |: ||  || || ::  ::
Db        140 --GSHSMSGWNSLHQASFQGNAETIKLLLKQGADRECQDDFGITPLFVAAQYGKLESMSI 197

Qy        665 LIKYNACVN--ATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTL------------KN 709
              ||   | ||  | ||   |||  |||:| |:   |||: |||| |
Db        198 LISSGANVNCQALDK--ATPLFIAAQEGHTKCVELLLSSGADPDLYCNEDNWQLPIHAAA 255

Qy        710 QEGQT-PLDLV----------SADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADA 758
              | | |  |||:          | || : :|        | :  : ||  || |
Db        256 QMGHTETLDLLIPRTNRACDTEPDKVSPVYSAVF--GGREECLEMLLQNGY-SPDAQM-C 311

Qy        759 LSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 795
              | | |||  ::     |     || : :::   | :
Db        312 LVFGFSSPLCMAFQKDCD-----FSGIVNILLKYGAQ 343


RESULT 27
V245_FOWPV
ID   V245_FOWPV     STANDARD;      PRT;   436 AA.
AC   Q9J4Z5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV245.
GN   FPV245.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44589.1; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 9.
DR   SMART; SM00248; ANK; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT        1     30       ANK 1.
FT   REPEAT       34     63       ANK 2.
FT   REPEAT       67     96       ANK 3.
FT   REPEAT       98    119       ANK 4.
FT   REPEAT      123    152       ANK 5.
FT   REPEAT      156    185       ANK 6.
FT   REPEAT      189    218       ANK 7.
FT   REPEAT      222    252       ANK 8.
FT   REPEAT      253    283       ANK 9.
FT   REPEAT      287    317       ANK 10.
SQ   SEQUENCE   436 AA;  49509 MW;  7F74EFE846B876C0 CRC64;

  Query Match             5.6%;  Score 321;  DB 1;  Length 436;
  Best Local Similarity   27.0%;  Pred. No. 3.9e-12;
  Matches  100;  Conservative   60;  Mismatches  148;  Indels   62;  Gaps    7;

Qy        318 ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC 377
              :|:  ||  ||   :|  |||  |     :  :||: :|| |:|| ::      |  |
Db         19 KLIKNKGNCINISVEETTTPLIDAIRTGNAKIVELFIKHGAQVNHVNTKIPNPLLTAIKI 78

Qy        378 GHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAG 437
              |     :|||  | | :|: :       :  | :: :|  |: :    |
```

```
Db         79 GSHDIVKLLLINGVDTSILPVP-----CINKEMIKTILDSGVKVNTKNA---------- 122

Qy        438 DVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPL 497

Db        123 --------------------KSKTFLHYAIKNNDLEVIKMLFEYGADVNIKDDNGCYPI 161

Qy        498 HNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNR 557

Db        162 HIATRSNSYEIIKLLLEKGAYANVKDNYGNSPLHNAAKYGDYACIKLVLDHTNNISNKCN 221

Qy        558 DGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPL 617

Db        222 NGVTPLH---------------NAILYNR------SAVELLINNRSINDTDVDG--YTPL 258

Qy        618 HLAAGYN-NLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD-VAALLIKYNACVNAT 675

Db        259 HYALQPPCSIDIIDILLYNNADISIKDNNGRNPIDTAFKYINRDSVIKELLANAVLINEV 318

Qy        676 DKWAFTPLHE 685

Db        319 GKLKDTTILE 328


RESULT 28
V246_FOWPV
ID   V246_FOWPV     STANDARD;      PRT;   592 AA.
AC   Q9J4Z4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative ankyrin-repeat protein FPV246.
GN   FPV246.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 10 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44590.1; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 3.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT       63     93       ANK 1.
FT   REPEAT      162    191       ANK 2.
FT   REPEAT      195    224       ANK 3.
FT   REPEAT      227    255       ANK 4.
FT   REPEAT      256    284       ANK 5.
FT   REPEAT      288    317       ANK 6.
FT   REPEAT      321    348       ANK 7.
FT   REPEAT      352    381       ANK 8.
FT   REPEAT      386    413       ANK 9.
FT   REPEAT      417    449       ANK 10.
SQ   SEQUENCE   592 AA;  66850 MW;  4529996B135848DC CRC64;
```

```
  Query Match             5.5%;  Score 319;  DB 1;  Length 592;
  Best Local Similarity   21.5%;  Pred. No. 8.2e-12;
  Matches  131;  Conservative   89;  Mismatches  213;  Indels  176;  Gaps   20;

Qy          5 KDVVEYLLQNGASVQARDDGGLI--------PLHNACSFGHAEVVNLLLRHG--ADPNAR 54

Db         36 RDRIEEIIELNRIIELDDDDVIYDQIYFPYNPLHQVIEARRANLVEIILDQGKYTVNSVN 95

Qy         55 DNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE 114

Db         96 DDFIFYPLH-------VLTCV-----------------------PETSDVLT--YL--- 119

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174

Db        120 --------NEIEEFTLVKDLEL--------RACKMNLSKSISIAIIKQVL----KGIREF 159

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234

Db        160 TDNDLRKLDNQIREDELKIAQLLISRGAKLDIKNEYGYTPLRNTVINGNFALTKLLLDNG 219

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294

Db        220 ADAS-IKCNGMTIFEISTLSQ--------------------NVDMIKEIIK------RC 251

Qy        295 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 354

Db        252 GYNHDSKILCRVAS--KGYINVIHFLLDIGFNVNSFDSFGETPLHAATRSGSVETVNALI 309

Qy        355 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 414

Db        310 SYGSTVDIKDNIGSTPLMHA--CGNKDISKLLIEKGADPNISNIHGY------------- 354

Qy        415 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 474

Db        355 ----------------------------------------------TPLHNAAAYGSVD 367

Qy        475 VVEYLLQHGADVHAKDK-GGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEA 533

Db        368 VVNLLLSYGVSIDAKDKIIGSTPLDHGVH--HPEIVKLLLEKGANPNTINLYGYTPLKNA 425

Qy        534 AAKGK----YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKK 589

Db        426 ILKSRLSAEYIIPYVVLQDFSILDGRNMPGS--------GIN--MELIRND-KLLQNIKT 474

Qy        590 GCLARVKKL 598

Db        475 SCEIELKRM 483


RESULT 29
ANS1_HUMAN
ID   ANS1_HUMAN     STANDARD;      PRT;  1134 AA.
AC   Q92625;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat and SAM domain containing protein 1.
GN   ANKS1 OR KIAA0229.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=97191544; PubMed=9039502;
RA   Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayasi Y.,
RA   Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT   "Prediction of the coding sequences of unidentified human genes. VI.
RT   The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT   analysis of cDNA clones from cell line KG-1 and brain.";
RL   DNA Res. 3:321-329(1996).
CC   -!- SIMILARITY: Contains 6 ANK repeats.
```

```
CC   -!- SIMILARITY: Contains 1 PID domain.
CC   -!- SIMILARITY: Contains 2 SAM domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D86982; BAA13218.1; ALT_INIT.
DR   HSSP; Q02410; 1AQC.
DR   Genew; HGNC:20961; ANKS1.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR006020; PTB_PID.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 7.
DR   Pfam; PF00640; PID; 1.
DR   Pfam; PF00536; SAM; 2.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00462; PTB; 1.
DR   SMART; SM00454; SAM; 2.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS01179; PID; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 2.
KW   ANK repeat; Repeat.
FT   DOMAIN        1     55       GLY-RICH.
FT   REPEAT       79    108       ANK 1.
FT   REPEAT      112    141       ANK 2.
FT   REPEAT      148    177       ANK 3.
FT   REPEAT      181    210       ANK 4.
FT   REPEAT      214    243       ANK 5.
FT   REPEAT      246    275       ANK 6.
FT   DOMAIN      696    762       SAM 1.
FT   DOMAIN      770    837       SAM 2.
FT   DOMAIN      936   1091       PID.
SQ   SEQUENCE   1134 AA;  123081 MW;  21B7171BCE4800BA CRC64;

  Query Match             5.5%;  Score 317;  DB 1;  Length 1134;
  Best Local Similarity   21.3%;  Pred. No. 2.9e-11;
  Matches  174;  Conservative  113;  Mismatches  278;  Indels  252;  Gaps   29;

Qy        265 KGHSLLQAAREADVTRIKKHLSLEMV--------------------------NFKHPQT 297
              |   ||:|||    :  ::| || : :                          :  ||
Db          3 KEQELLEAARTGHLPAVEKLLSGKRLSSGFGGGGGGGSGGGGGGSGGGGGGLGSSSHP-- 60

Qy        298 HETALHCAAASPYPKRKQICELL-LRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKH 356
                              :  || : :| |:|       ||||  |:   | |||||:::: 
Db         61 ------------------LSSLLSMWRGPNVNCVDSTGYTPLHHAALNGHKDVVEVLLRN 102

Qy        357 EAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG-------FTALQMGNE 409
              :|  |  |: |    || ||: |  |  |||:  |      :: |           | | |: 
Db        103 DALTNVADSKGCYPLHLAAWKGDAQIVRLLIHQGPSHTRVNEQNNDNETALHCAAQYGHT 162

Qy        410 NVQQLLQEGI---SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHF 466
                | ::| | :   :: |:: :   |  ||  | :| || |       |       :: ||||
Db        163 EVVKVLLEELTDPTMRNNKFETPLDLAALYGRLEVVKMLLNAHP-NLLSCNTKKHTPLHL 221

Qy        467 AAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK 526
              ||      :||: ||  | | : : :  |  || | :|  :| ::|:  |  ||: |
Db        222 AARNGHKAVVQVLLDAGMDSNYQTEMGSA-LHEAALFGKTDVVQILLAAGTDVNIKDNHG 280

Qy        527 FTPL---HEAAAKGKYEICKLLLQH--GADPTK---KNRDGNTPL--------------- 563
                |  |    |  ::   :|  |: |  |   ||   |       ||
Db        281 LTALDTVRELPSQKSQQIAALIEDHMTGKRSTKEVDKTPPPQPPLISSMDSISQKSQGDV 340

Qy        564 -----DLVKDGDTDIQD-----------------LLRGDAALLDAAKKGCLARVKKLSS 600
                   :|: | | : ::                 :  | ::  |:  |   |   |
Db        341 EKAVTELIIDFDANAEEEGPYEALYNAISCHSLDSMASGRSSDQDSTNKEAEAAGVK--- 397

Qy        601 PDNVNCRDTQGRHSTPL-------HLAAGYNNLEVAEYLLQ----HGADVNAQDKGGLIP 649
              | | |:     : |         |:   | |  :| |       ||:    :|:
Db        398 PAGVRPRERPPPPAKPPPDEEEEDHIDKKYFPLTASEVLSMRPRIHGSAAREEDE----- 452

Qy        650 LHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKN 709
                    |   | :   |   |  | | :  |:|      | | |  :
Db        453 --------HPYELLLTAETKKVVLVDGK---TKDHRRSSSSRSQDS----AEGQDGQVPE 497

Qy        710 Q-----EGQTPLDLVSADDVSALLTAAMP-PSALP---SCYKPQVL---NGVRSPGAT-- 755
              |     | :|:  |  |  |  | ||    |   |  :|:|  :     || :|||:
Db        498 QFSGLLHGSSPVCEVGQDPFQLLCTAGQSHPDGSPQQGACHKASMQLEETGVHAPGASQP 557

Qy        756 ------------------------------------------------------------ 755

Db        558 SALDQSKRVGYLTGLPTTNSRSHPETLTHTASPHPGGAEEGDRSGARSRAPPTSKPKAEL 617

Qy        756 ----------ADALSSGPSSPSSL-SAASSLDNLS-------------GSFSELSSVVSS 791
                         :| |:  |:  ::: | : || | |                : |:  ::||
Db        618 KLSRSLSKSDSDLLTCSPTEDATMGSRSESLSNCSIGKKRLEKSPSFASEWDEIEKIMSS 677

Qy        792 SGTEGASSLEKKEVPG---VDFSITQFVRNLGLEH-----LMDIFEREQITLDVLVEMGH 843
                |       | |:::: |    :: |: ::: ::||:       |:: |:        ::|
Db        678 IGEGIDFSQERQKISGSRTLEQSVGEWLESIGLQQYESKLLLNGFDDVHFLGSNVME--E 735

Qy        844 KELKEIGINAYGHRHKLIKGVERL----ISGQQGLNP 876
              ::|::|||:   || ||::    |       |  | :|
Db        736 QDLRDIGISDPQHRRKLLQAARSLPKVKALGYDGNSP 772


RESULT 30
PPGA_HUMAN
ID   PPGA_HUMAN     STANDARD;      PRT;   528 AA.
AC   Q96I34;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16A.
GN   PPP1R16A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RT   "The nucleotide sequence of a long cDNA clone isolated from human
RT   spleen.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung, Placenta, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
CC   -!- FUNCTION: Inhibits protein phosphatase 1 activity toward
CC       phosphorylase, myosin light chain and myosin substrates (By
CC       similarity).
CC   -!- SUBUNIT: Binds PP1.
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (By
CC       similarity).
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   -!- CAUTION: Ref.1 sequence may be incorrect due to a cloning
CC       artifact.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK090471; BAC03452.1; ALT_SEQ.
DR   EMBL; BC007854; AAH07854.1; -.
DR   EMBL; BC049841; AAH49841.1; -.
DR   EMBL; BC053506; AAH53506.1; -.
DR   Genew; HGNC:14941; PPP1R16A.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 4.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   DOMAIN       18     45       COILED COIL (POTENTIAL).
FT   REPEAT       70     99       ANK 1.
FT   REPEAT      103    132       ANK 2.
FT   REPEAT      136    165       ANK 3.
FT   REPEAT      231    260       ANK 4.
FT   REPEAT      264    293       ANK 5.
FT   LIPID       525    525       S-farnesyl cysteine (Potential).
FT   CONFLICT     34     71       MISSING (IN REF. 1).
FT   CONFLICT    249    249       A -> P (IN REF. 1).
FT   CONFLICT    272    272       A -> P (IN REF. 1).
SQ   SEQUENCE   528 AA;  57811 MW;  A5A2D5E1A16A438C CRC64;

  Query Match             5.4%;  Score 310;  DB 1;  Length 528;
  Best Local Similarity   26.2%;  Pred. No. 2.4e-11;
  Matches  116;  Conservative   55;  Mismatches  164;  Indels  108;  Gaps   10;

Qy        413 QLLQEGISLGNSEADRQLLEA------------------AKAGDVETVKKLCTVQSVNC 453
              :|| |   :|      :| |                    | |  |  :|    | :
Db          6 ELLAEMPMVGRMSTQERLKHAQKRRAQQVKMWAQAEKEAQGKKGPGERPRKEAASQGL-L 64

Qy        454 RDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLV 513
              : :   |  |  ||  | :  |   |  |      :: ||  ||  |      |: : |:
Db         65 KQVLFPPSVVLLEAAARNDLEEVRQFLGSGVSPDLANEDGLTALHQCCIDDFREMVQQLL 124

Qy        514 KHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT-D 572
              : || :|  |   :|||| ||  |   : :||:  ||:    | ||| | ||  |  | |
Db        125 EAGANINACDSECWTPLHAAATCGHLHLVELLIASGANLLAVNTDGNMPYDLCDDEQTLD 184

Qy        573 IQDLLRGDAAL----LDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 628
                :   |  :    ::||:      | :|   |::  |
Db        185 CLETAMADRGITQDSIEAAR-----AVPELRMLDDIRSR---------------------- 218

Qy        629 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQ 688
              || |||::|    |   || ||: |  : ||||::: | ::| |:  : ||| ||
Db        219 ----LQAGADLHAPLDHGATLLHVAAANGFSEAAALLLEHRASLSAKDQDGWEPLHAAAY 274

Qy        689 KGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNG 748
               |:  |  ||:|||||   |:   :||||:   ::| | |        |   :|
Db        275 WGQVPLVELLVAHGADLNAKSLMDETPLDVCGDEEVRAKL--------LELKHK------ 320

Qy        749 VRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGV 808
                  |||    |  ||                |    ||:|: |
Db        321 -------HDALLRAQSRQRSL--------------LRRRTSSAGSRG------------ 346

Qy        809 DFSITQFVRNLGLEHLMDIFERE 831
                  : || : |    |:: ::
Db        347 -----KVVRRVSLTQRTDLYRKQ 364


RESULT 31
PPGA_MOUSE
ID   PPGA_MOUSE     STANDARD;      PRT;   524 AA.
AC   Q923M0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16A (Myosin
DE   phosphatase targeting subunit 3) (Mypt3).
GN   PPP1R16A OR MYPT3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND FARNESYLATION.
RX   MEDLINE=21234869; PubMed=11336659;
RA   Skinner J.A., Saltiel A.R.;
RT   "Cloning and identification of MYPT3: a prenylatable myosin targetting
RT   subunit of protein phosphatase 1.";
RL   Biochem. J. 356:257-267(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Inhibits protein phosphatase 1 activity toward
CC       phosphorylase, myosin light chain and myosin substrates.
CC   -!- SUBUNIT: Binds PP1.
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (By
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Highly expressed in testis, followed by liver,
CC       heart, brain and lung.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY010723; AAG40949.1; -.
DR   EMBL; BC057450; AAH57450.1; -.
DR   MGD; MGI:1920312; Ppp1r16a.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
```

```
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   REPEAT       70     99       ANK 1.
FT   REPEAT      103    132       ANK 2.
FT   REPEAT      136    165       ANK 3.
FT   REPEAT      231    260       ANK 4.
FT   REPEAT      264    293       ANK 5.
FT   DOMAIN      448    469       COILED COIL (POTENTIAL).
FT   LIPID       521    521       S-farnesyl cysteine (Probable).
SQ   SEQUENCE   524 AA;  57529 MW;  A0F682A794475429 CRC64;

  Query Match             5.4%;  Score 309;  DB 1;  Length 524;
  Best Local Similarity   29.4%;  Pred. No. 2.7e-11;
  Matches  109;  Conservative   42;  Mismatches  142;  Indels   78;  Gaps    9;

Qy        461 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 520
Db         72 SVALLEAAARNDLEEVRQFLTSGVSPNLANEDGLTALHQCCIDDFQEMAQQLLDAGADVN 131

Qy        521 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 580
Db        132 ARDSECWTPLHAAATCGHLHLVELLISRGADLLAVNSDGNMPYDLCEDAQT--LDCLETA 189

Qy        581 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 640
Db        190 MANQGITQEGI---EEARAVPELCMLNDLQNR------LQAGANLSDP----LDHGATL- 235

Qy        641 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 700
Db        236 ---------LHIAAANGFSEVATLLLEQGASLSAKDHDGWEPLHAAAYWGQVHLVELLVA 286

Qy        701 HGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS 760
Db        287 HGADLNGKSLVDETPLDVCGDEEVRAKL--------LELKHK-------------QDALL 325

Qy        761 SGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLG 820
Db        326 RAQGRQRSL---------------LRRRTSSAGSRG-----------------KVVRRVS 353

Qy        821 LEHLMDIFERE 831
Db        354 LTHRTNLYRKE 364


RESULT 32
PPGB_MOUSE
ID   PPGB_MOUSE     STANDARD;      PRT;   568 AA.
AC   Q8VHQ3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
DE   inhibited membrane-associated protein) (CAAX box protein TIMAP).
GN   PPP1R16B.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Cerebellum;
RX   MEDLINE=22525552; PubMed=12638124;
RA   Magdaleno S., Northcutt G.M., Curran T., Kurschner C.;
RT   "mPPP1R16B is a novel mouse protein phosphatase 1 targeting subunit
RT   whose mRNA is located in cell bodies and dendrites of neurons in four
RT   distinct regions of the brain.";
RL   Gene Expr. Patterns 1:143-149(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be a downstream target for TGF-beta1 signaling
CC       cascade in endothelial cells (By similarity).
CC   -!- SUBUNIT: Binds PP1 (Probable).
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (By
CC       similarity).
CC   -!- MISCELLANEOUS: Repressed by TGF-beta1 which requires denovo
CC       protein synthesis and histone deacetylase (By similarity).
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF423761; AAL62093.1; -.
DR   EMBL; BC054764; AAH54764.1; -.
DR   EMBL; BC057542; AAH57542.1; -.
DR   MGD; MGI:2151841; Ppp1r16b.
DR   GO; GO:0005886; C:plasma membrane; ISS.
DR   GO; GO:0019903; F:protein phosphatase binding; ISS.
DR   GO; GO:0007165; P:signal transduction; ISS.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   DOMAIN       15     55       COILED COIL (POTENTIAL).
FT   REPEAT      100    129       ANK 1.
FT   REPEAT      133    162       ANK 2.
FT   REPEAT      228    257       ANK 3.
FT   REPEAT      261    290       ANK 4.
FT   REPEAT      531    560       ANK 5.
FT   LIPID       565    565       S-farnesyl cysteine (Potential).
SQ   SEQUENCE   568 AA;  63571 MW;  5D8446C81DEBA3D1 CRC64;

  Query Match             5.4%;  Score 309;  DB 1;  Length 568;
  Best Local Similarity   30.0%;  Pred. No. 3.1e-11;
  Matches   89;  Conservative   53;  Mismatches  121;  Indels   34;  Gaps    9;

Qy          8 VEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAI 67
Db         84 VRYFLKNKVSPDLCNEDGLTALHQCCIDNFEEIVKLLLSHGANVNAKDNELWTPLHAAAT 143

Qy         68 KGKIDVCIVLLQHGAEPTIRNTDGRTALDLA--DPSAKAVLT-GEYK--KDELLESARSG 122
Db        144 CGHINLVKILVQYGADLLAVNSDGNMPYDLCEDEPTLDVIETCMAYQGITQEKINEMRAA 203
```

```
Qy        123 NEEKMMALLTPLNVNCHAS--------DGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
Db        204 PEQKMIS-----DIHCMIAAGQDLDWIDGQGATLLHIAGANGYLRAAELLLDHGVRVDVK 258

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
Db        259 DWDGWEPLHAAAFWGQMPMAELLVSHGASLSARTSMDEMPI-DLCEEEEFKVLLLELKHK 317

Qy        235 ADPTLLN-CHNKSAID----------LAPTPQLKER---LAYEFKGHSLLQAAREA 276
Db        318 HDVIMKSQLRHKSSLSRRTSSAGSRGKVVRRASLSDRTNLYRKEYEGEAILWQQRSA 374


RESULT 33
ASB3_HUMAN
ID   ASB3_HUMAN     STANDARD;      PRT;   518 AA.
AC   Q9Y575; Q9NVZ2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ankyrin repeat and SOCS box containing protein 3 (ASB-3).
GN   ASB3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20564172; PubMed=11111040;
RA   Kile B.T., Viney E.M., Willson T.A., Brodnicki T.C., Cancilla M.R.,
RA   Herlihy A.S., Croker B.A., Baca M., Nicola N.A., Hilton D.J.,
RA   Alexander W.S.;
RT   "Cloning and characterization of the genes encoding the ankyrin repeat
RT   and SOCS box-containing proteins Asb-1, Asb-2, Asb-3 and Asb-4.";
RL   Gene 258:31-41(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 11 ANK repeats.
CC   -!- SIMILARITY: Contains 1 SOCS box domain.

CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF156778; AAD41895.1; -.
DR   EMBL; AK000985; BAA91455.1; -.
DR   EMBL; AK001283; BAA91599.1; -.
DR   EMBL; BC006488; AAH06488.1; -.
DR   EMBL; BC009569; AAH09569.1; -.
DR   Genew; HGNC:16013; ASB3.
DR   MIM; 605760; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001496; SOCS_C.
DR   Pfam; PF00023; ank; 10.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 10.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50225; SOCS; 1.
KW   ANK repeat; Repeat.
FT   REPEAT       9     38       ANK 1.
FT   REPEAT      42     71       ANK 2.
FT   REPEAT      78    107       ANK 3.
FT   REPEAT     111    140       ANK 4.
FT   REPEAT     145    174       ANK 5.
FT   REPEAT     178    207       ANK 6.
FT   REPEAT     211    240       ANK 7.
FT   REPEAT     246    275       ANK 8.
FT   REPEAT     279    308       ANK 9.
FT   REPEAT     315    346       ANK 10.
FT   REPEAT     348    373       ANK 11.
FT   DOMAIN     441    504       SOCS BOX.
FT   CONFLICT   118    118       L -> S (IN REF. 2; BAA91599).
SQ   SEQUENCE   518 AA;  57745 MW;  E6FC2EF76DD814DF CRC64;

  Query Match             5.3%;  Score 306.5;  DB 1;  Length 518;
  Best Local Similarity   30.2%;  Pred. No. 3.8e-11;
  Matches  111;  Conservative   45;  Mismatches  113;  Indels   99;  Gaps   14;

Qy        433 AAKAGDVETVKKLCTV-QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGAD---VHA 488
Db         17 AAREGNVKVLRKLLKKGRSVDVADNRGWM--PIHEAAYHNSVECLQMLINADSSENYIKM 74

Qy        489 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 548
Db         75 KTFEGFCALHLAASQGHWKIVQILLEAGADPNATTLEETTPLFLAVENGQIDVLRLLLQH 134

Qy        549 GADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRD 608
Db        135 GA----------------------------------------------------NVN--- 139

Qy        609 TQGRHS----TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAAL 664
Db        140 --GSHSMCGWNSLHQASFQENAEIIKLLLRKGANKECQDDFGITPLFVAAQYGKLESLSI 197

Qy        665 LIKYNACVN--ATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTL------------KN 709
Db        198 LISSGANVNCQALDK--ATPLFIAAQEGHTKCVELLLSSGADPDLYCNEDSWQLPIHAAA 255

Qy        710 QEGQTP-LDLV----------SADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADA 758
Db        256 QMGHTKILDLLIPLTNRACDTGLNKVSPVYSAVF--GGHEDCLEILLRNGY-SPDAQA-C 311

Qy        759 LSSGPSSP 766
Db        312 LVFGFSSP 319
```

```
RESULT 34
RN5A_MOUSE
ID   RN5A_MOUSE      STANDARD;      PRT;   735 AA.
AC   Q05921; Q9ERU7;
DT   01-NOV-1995 (Rel. 32, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN   RNASEL OR RNS4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C3H/An; TISSUE=Adipose tissue;
RX   MEDLINE=20515260; PubMed=11063255;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes: mediators
RT   of interferon action.";
RL   Mamm. Genome 11:989-992(2000).
RN   [2]
RP   SEQUENCE OF 1-679 FROM N.A.
RX   MEDLINE=93201598; PubMed=7680958;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAase: a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:753-765(1993).
RN   [3]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11585831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";
RL   J. Biol. Chem. 276:48473-48482(2001).
RN   [4]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13354-13354(2002).
CC   -!- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of mRNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 3' of UpNp dimers, with preference for
CC       UU and UA sequences, to sets of discrete products ranging from
CC       between 4 and 22 nucleotides in length.
CC   -!- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC       either manganese or magnesium and ATP.
CC   -!- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC       linked oligoadenylates) the homodimerization and subsequent
CC       activation occurs. Inhibited by RNase L inhibitor.
CC   -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC   -!- TISSUE SPECIFICITY: Expressed in spleen, thymus, lung, testis,
CC       kidney, liver and heart.
CC   -!- INDUCTION: By interferons.
CC   -!- DOMAIN: A single active nuclease domain in a dimer is
CC       sufficient for ribonuclease activity.
CC   -!- DOMAIN: The kinase domain allows the homodimerization.
CC   -!- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC       constitute the 2-5A binding domain.
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF281045; AAG33708.1; -.
DR   EMBL; L10382; AAA37117.1; -.
DR   PIR; B45771; B45771.
DR   HSSP; P42773; 1IHB.
DR   MGD; MGI:1098272; Rnasel.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR006567; PUG.
DR   Pfam; PF00023; ank; 8.
DR   Pfam; PF00069; pkinase; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00580; PUG; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger;
KW   Repeat; ANK repeat.
FT   REPEAT      24     53       ANK 1.
FT   REPEAT      58     87       ANK 2.
FT   REPEAT      91    120       ANK 3.
FT   REPEAT     124    153       ANK 4.
FT   REPEAT     167    197       ANK 5.
FT   REPEAT     201    234       ANK 6.
FT   REPEAT     238    268       ANK 7.
FT   REPEAT     272    301       ANK 8.
FT   REPEAT     303    328       ANK 9.
FT   DOMAIN     364    584       PROTEIN KINASE LIKE.
FT   DOMAIN     585    735       RIBONUCLEASE.
FT   DOMAIN     229    242       2-5A BINDING (P-LOOP) 1.
FT   DOMAIN     253    275       2-5A BINDING (P-LOOP) 2.
FT   ZN_FING    401    436       C6-TYPE (POTENTIAL).
SQ   SEQUENCE   735 AA;  83274 MW;  B6632F4A5B50F711 CRC64;

  Query Match             5.3%;  Score 305;  DB 1;  Length 735;
  Best Local Similarity   27.5%;  Pred. No. 7.9e-11;
  Matches  148;  Conservative   69;  Mismatches  232;  Indels   90;  Gaps   23;

Qy        425 EADRQLLEAAKAGDVETVKKLCT--VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 482
              | |  |::| : |||  |::|      :  | |  |  ||||  |    || :|  || |
Db         24 EDDSSLIKAVQKGDVVRVQQLLEKGADANACEDTWG--WTPLHNAVQAGRVDIVNLLLSH 81

Qy        483 GADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEIC 542
              ||| | : | |  |    |   | :: |:|:  || ||  |   ||   ||| :|  |
Db         82 GADPHRRKKNGATPFIIAGIQGDVKLLEILLSCGADVNECDENGFTAFMEAAERGNAEAL 141

Qy        543 KLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDA-ALLDAAKKGCL--ARVKKLS 599
              : |  ||:           ::| :   | : | :| | ||: ||:|| |   |:
Db        142 RFLFAKGAN----------VNLRRQTTKDKRRLKQGGATALMSAAEKGHLEVLRILLND 190

Qy        600 SPDNVNCRDTQGRHS---TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 656
              |: ||  ||::   | |:      |  |:   |:|||||| : :  |  ||  |
Db        191 MKAEVDARDNMGRNALIRTLLNWDC-ENVEEITSILIQHGADVNVRGERGKTPLIAAVER 249

Qy        657 GHVDVAALLIKYNAC-VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 715
              |  :  :|:      ::| |   | |  |  |   ::  |||  |||
Db        250 KHTGLVQMLLSREGINIDARDNEGKTALLIAVDKQLKEIVQLLLEKGAD---------- 298

Qy        716 LDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSL 775
                ||:   :        |    | | |   | |       |  || |   ::|
Db        299 ----KCDDL-VWIARRNHDYHLVKLLLPYVANPDTDPPA-------GDWSPHSSRWGTAL 346

Qy        776 DNLSG----SFSELSSVVSS----SGT-EGASSL---EKKEVPGVDF--------SITQF 815
               :|       :|  :      :|| |||  |   : :||     |
Db        347 KSLHSMTRPMIGKLKIFIHDDYKIAGTSEGAVYLGIYDNREVAVKVFRENSPRGCKEVSC 406

Qy        816 VRNLGLEH--LMDIFERE------------QITLDVLVEMGHKELKEIGINAYGHR--H 858
              :|: | :|  |:  : ||             : ||:  : :  :|  | | : : |
Db        407 LRDCG-DHSNLVAFYGREDDKGCLYVCVSLCEWTLEEFLRLPREEPVENGEDKFAHSILL 465
```

```
Qy        859 KLIKGVERL-ISG--QQGLNPYLTLNTSGSGTILIDLSPDDK---EFQSVEEEMQSTVR 911
Db        466 SIFEGVQKLHLHGYSHQDLQPQNILIDSKKAVRLADFDQSIRWMGESQMVRRDLEDLGR 524

RESULT 35
PPGB_BOVIN
ID   PPGB_BOVIN     STANDARD;      PRT;   568 AA.
AC   Q95N27;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
DE   inhibited membrane-associated protein) (bTIMAP) (CAAX box protein
DE   TIMAP).
GN   PPP1R16B.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22049828; PubMed=12055102;
RA   Cao W., Mattagajasingh S.N., Xu H., Kim K., Fierlbeck W., Deng J.,
RA   Lowenstein C.J., Ballermann B.J.;
RT   "TIMAP, a novel CAAX box protein regulated by TGF-beta1 and expressed
RT   in endothelial cells.";
RL   Am. J. Physiol. 283:C327-C337(2002).
CC   -!- FUNCTION: May be a downstream target for TGF-beta1 signaling
CC       cascade in endothelial cells.
CC   -!- SUBUNIT: Binds PP1 (Probable).
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (Potential).
CC   -!- INDUCTION: Inhibited by TGF-beta1.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF362909; AAK52795.1; -.
DR   GO; GO:0005886; C:plasma membrane; ISS.
DR   GO; GO:0019903; F:protein phosphatase binding; ISS.
DR   GO; GO:0007165; P:signal transduction; ISS.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   DOMAIN       15     55       COILED COIL (POTENTIAL).
FT   REPEAT      100    129       ANK 1.
FT   REPEAT      133    162       ANK 2.
FT   REPEAT      228    257       ANK 3.
FT   REPEAT      261    290       ANK 4.
FT   REPEAT      531    560       ANK 5.
FT   LIPID       565    565       S-farnesyl cysteine (Potential).
SQ   SEQUENCE   568 AA;  63643 MW;  61C10899462B0334 CRC64;

  Query Match             5.3%;  Score 303.5;  DB 1;  Length 568;
  Best Local Similarity   25.1%;  Pred. No. 6.6e-11;
  Matches  102;  Conservative   49;  Mismatches  124;  Indels  131;  Gaps   11;

Qy        425 EADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGA 484
Db         67 EASVALLEASLRNDAEEVR----------------------------------YFLKNKV 92

Qy        485 DVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKL 544
Db         93 SPDLCNEDGLTALHQCCIDNFEEIVKLLLSHGANVNAKDNELWTPLHAAATCGHINLVKI 152

Qy        545 LLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLA-------RVKK 597
Db        153 LVQYGADLLAVNSDGNMPYDLCEDEPT------------LDVIET-CMAYQGITQEKINE 199

Qy        598 L-SSPD-----NVNCR----------DTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNA 641
Db        200 MRAAPEQQMISDIHCMIAAGQDLDWVDAQG--ATLLHIAGANGYLRAAELLLDHGVRVDV 257

Qy        642 QDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAH 701
Db        258 KDWDGWEPLHAAAFWGQM--------------------------------QMAELLVSH 284

Qy        702 GADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSS 761
Db        285 GASLSARTSMDEMPIDLCEEEEFKVLL--------LELKHKHDVI-------------- 321

Qy        762 GPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPG 807
Db        322 ---MKSQLRHKSSLSRRTSSAGSRGKVVRRASLSDRTNLYRKEYEG 364

RESULT 36
PPGB_HUMAN
ID   PPGB_HUMAN     STANDARD;      PRT;   567 AA.
AC   Q96T49; O94912; Q9NQG4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein phosphatase 1 regulatory inhibitor subunit 16B (TGF-beta-
DE   inhibited membrane-associated protein) (hTIMAP) (CAAX box protein
DE   TIMAP) (Ankyrin repeat domain protein 4).
GN   PPP1R16B OR ANKRD4 OR KIAA0823.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22049828; PubMed=12055102;
RA   Cao W., Mattagajasingh S.N., Xu H., Kim K., Fierlbeck W., Deng J.,
RA   Lowenstein C.J., Ballermann B.J.;
RT   "TIMAP, a novel CAAX box protein regulated by TGF-beta1 and expressed
RT   in endothelial cells.";
RL   Am. J. Physiol. 283:C327-C337(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99156230; PubMed=10048485;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 5:355-364(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
```

```
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
CC   -!- FUNCTION: May be a downstream target for TGF-beta1 signaling
CC       cascade in endothelial cells.
CC   -!- SUBUNIT: Binds PP1 (Probable).
CC   -!- SUBCELLULAR LOCATION: Attached to the plasma membrane (Potential).
CC   -!- TISSUE SPECIFICITY: Highly expressed in vascular endothelium, CNS,
CC       lung, spleen, kidney and testis.
CC   -!- INDUCTION: Inhibited by TGF-beta1.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF362910; AAK52796.1; -.
DR   EMBL; AB020630; BAA74846.2; ALT_INIT.
DR   EMBL; AL121889; CAB98284.1; -.
DR   EMBL; AL031657; -; NOT_ANNOTATED_CDS.
DR   Genew; HGNC:15850; PPP1R16B.
DR   GO; GO:0005886; C:plasma membrane; NAS.
DR   GO; GO:0019903; F:protein phosphatase binding; NAS.
DR   GO; GO:0007165; P:signal transduction; NAS.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 5.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat; Coiled coil; Lipoprotein; Prenylation; Membrane.
FT   DOMAIN       15     55       COILED COIL (POTENTIAL).
FT   REPEAT      100    129       ANK 1.
FT   REPEAT      133    162       ANK 2.
FT   REPEAT      228    257       ANK 3.
FT   REPEAT      261    290       ANK 4.
FT   REPEAT      530    559       ANK 5.
FT   LIPID       564    564       S-farnesyl cysteine (Potential).
SQ   SEQUENCE   567 AA;  63551 MW;  8418CE26D0035CF1 CRC64;

  Query Match             5.2%;  Score 301.5;  DB 1;  Length 567;
  Best Local Similarity   25.1%;  Pred. No. 8.8e-11;
  Matches  102;  Conservative   48;  Mismatches  125;  Indels  131;  Gaps   11;

Qy        425 EADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGA 484
              ||   ||||:   | | |:                                  | |::
Db         67 EASVALLEASLRNDAEEVR---------------------------------YFLKNKV 92

Qy        485 DVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKL 544
                :: ||  ||  |      |: :||: ||| ||  |   :|||| ||  |   : |:
Db         93 SPDLCNEDGLTALHQCCIDNFEEIVKLLLSHGANVNAKDNELWTPLHAAATCGHINLVKI 152

Qy        545 LLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLA-------RVKK 597
              |:|:|||    | ||| | || :|  |            ||  :  |:|       :: :
Db        153 LVQYGADLLAVNSDGNMPYDLCEDEPT------------LDVIET-CMAYQGITQEKINE 199

Qy        598 LS-SPD-----NVNCR----------DTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNA 641
              :  :|:     :::|           | ||  :| ||:|     |  || || ||  |:

Db        200 MRVAPEQQMIADIHCMIAAGQDLDWIDAQG--ATLLHIAGANGYLRAAELLLDHGVRVDV 257

Qy        642 QDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAH 701
              :|  |  ||| || :| :                                  |:  ||::|
Db        258 KDWDGWEPLHAAAFWGQM---------------------------------QMAELLVSH 284

Qy        702 GADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSS 761
              ||  : :    : |:||   ::   ||        |  :|  |:
Db        285 GASLSARTSMDEMPIDLCEEEEFKVLL--------LELKHKHDVI--------------- 321

Qy        762 GPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPG 807
                | |   |||   : |       ||  :      ::| :||  |
Db        322 ---MKSQLRHKSSLSRRTSSAGSRGKVVRRASLSDRTNLYRKEYEG 364


RESULT 37
ANR6_HUMAN
ID   ANR6_HUMAN     STANDARD;      PRT;   692 AA.
AC   Q9Y2G4; Q9NU24; Q9UFQ9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ankyrin repeat domain protein 6.
GN   ANKRD6 OR KIAA0957.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [2]
RP   SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).
RC   TISSUE=Testis;
RA   Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 140-692 FROM N.A. (ISOFORM 1).
RA   Tracey A.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9Y2G4-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9Y2G4-2; Sequence=VSP_000270;
CC         Note=No experimental confirmation available;
CC   -!- SIMILARITY: Contains 8 ANK repeats.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AB023174; BAA76801.1; -.
DR   EMBL; AL117504; CAB55968.1; -.
DR   EMBL; AL096678; CAB86658.1; -.
DR   HSSP; Q00420; 1AWC.
DR   Genew; HGNC:17280; ANKRD6.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 8.
```

```
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Repeat; ANK repeat; Alternative splicing.
FT   REPEAT        9     38       ANK 1.
FT   REPEAT       41     70       ANK 2.
FT   REPEAT       74    103       ANK 3.
FT   REPEAT      107    136       ANK 4.
FT   REPEAT      140    169       ANK 5.
FT   REPEAT      173    202       ANK 6.
FT   REPEAT      206    235       ANK 7.
FT   REPEAT      239    268       ANK 8.
FT   VARSPLIC    264    264       Q -> QVLRFSRGRSLRKKRERLKEERRAQSVPRDEVAQSK
FT                                (in isoform 2).
FT                                /FTId=VSP_000270.
FT   CONFLICT     87     87       N -> I (IN REF. 2).
SQ   SEQUENCE   692 AA;  75722 MW;  D32115487A80839F CRC64;

  Query Match             5.2%;  Score 299.5;  DB 1;  Length 692;
  Best Local Similarity   21.1%;  Pred. No. 1.6e-10;
  Matches  175;  Conservative   86;  Mismatches  279;  Indels  291;  Gaps   28;

Qy         97 LADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTP---LNVNCHASDGRKSTPLHLAA 153
              ::   | | |:      : || :|  |  | :: |:      : | |   ||  |||||||
Db          1 MSQQDAVAALS-----ERLLVAAYKGQTENVVQLINKGARVAVTKH---GR--TPLHLAA  50

Qy        154 GYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFT 213
               : :||:||: | |:  :| ||   || |   |: |:   |: |  ::  |     |
Db         51 NKGHLPVVQILLKAGCDLDVQDDGDQTALHRATVVGNTEIIAALIHEGCALDRQDKDGNT 110

Qy        214 PLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAA 273
              ||||:                                          |   |:|
Db        111 ALHEAS-----------------------------------------WHGFSQSA 124

Qy        274 READVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKE 333
              :                                          ||:: |||:  | |
Db        125 K-------------------------------------------LLVKAGANVLAKNKA 140

Qy        334 FLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDP 393
              | ||:| : :|:    |::   :: :  :| | | || ||   ||   ||||:  |
Db        141 GNTALHLACQNSHSQSTRVLLLAGSRADLKNNAGDTCLHVAARYNHLSIIRLLLTAFCSV 200

Qy        394 NIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 453
              :                                    | :|||
Db        201 H-----------------------------------EKNQAGD--------------- 208

Qy        454 RDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLV 513
                  | || ||  |   | : ||: |||    :  |  ||  |  : : ||| ||
Db        209 --------TALHVAAALNHKKVAKILLEAGADTTIVNNAGQTPLETARYHNNPEVALLLT 260

Qy        514 K--HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN-------------RD 558
              |   |: |:  |    ||  | |   | |  :  |    :|  |:               |
Db        261 KAPQGS-VSAGD----TPSSEQAVARKEEAREEFLSASPEPRAKDDRRRKSRPKVSAFSD 315

Qy        559 GNTPLDLVKDGDTDIQ---------------------------DLLRG-DAALLDAAKK 589
              | |        ::                             | || |  :: |
Db        316 PTPPADQQPGHQKNLHAHNHPKKRNRHRCSSPPPPHEFRAYQLYTLYRGKDGKVMQAPIN 375

Qy        590 GCLAR--VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLE--VAEYLLQHGADVNAQDKG 645
              ||     : ||                 | ||  | |   :  |   : |||
Db        376 GCRCEPLINKLE----------------------NQLEATVEEIKAELG---SVQDKM 408

Qy        646 G--LIPLHNAASY-----GHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQL---- 694
              |  : |   :       : | |  :  |:|   : : |  ||  :| :  |
Db        409 NTKLGQMENKTQHQMRVLDKLMVERLSAERTECLNRLQQHSDTEKHE-GEKRQISLVDEL 467

Qy        695 ---CALLL------------AHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPS 739
                 | | :             |  |  |     |  | :| :        |  : : ||   |
Db        468 KTWCMLKIQNLEQKLSGDSRACRAKSTPSTCESSTGVDQLVVTAGPAAASDSSPPVVRP- 526

Qy        740 CYKPQVLNGVRSPGATADALSSGPSSPSSL------------SAASSLDNLSGSFSELSS 787
                | : ||   :     : |||   : | |             : |:  |  :|      :
Db        527 --KEKALNSTATQRLQQE-LSSSDCTGSRLRNVKVQTALLPMNEAARSDQQAGPCVNRGT 583

Qy        788 VVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLE----HLMDIFEREQIT 834
                ||  | :   : : |   :  |     |  |    |: |  :  ::|
Db        584 QTKKSGKSGPTR-HRAQQPAASSTCGQPPPATGSEQTGPHIRDTSQALELT 633


RESULT 38
Y045_METMA
ID   Y045_METMA     STANDARD;      PRT;   360 AA.
AC   Q8Q0U0;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical ANK-repeat protein MM0045.
GN   MM0045.
OS   Methanosarcina mazei (Methanosarcina frisia).
OC   Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC   Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2209;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Goe1 / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX   MEDLINE=22120827; PubMed=12125824;
RA   Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA   Martinez-Arias R., Henne A., Wiezer A., Baeumer S., Jacobi C.,
RA   Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA   Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA   Fritz H.-J., Gottschalk G.;
RT   "The genome of Methanosarcina mazei: evidence for lateral gene
RT   transfer between Bacteria and Archaea.";
RL   J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE013224; AAM29741.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PROSITE; PS50088; ANK_REPEAT; 7.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat; Complete proteome.
FT   REPEAT      100    129       ANK 1.
FT   REPEAT      133    162       ANK 2.
FT   REPEAT      166    195       ANK 3.
FT   REPEAT      199    228       ANK 4.
FT   REPEAT      232    261       ANK 5.
FT   REPEAT      265    294       ANK 6.
FT   REPEAT      298    327       ANK 7.
SQ   SEQUENCE   360 AA;  39044 MW;  C20B57832367A592 CRC64;

  Query Match             5.1%;  Score 295.5;  DB 1;  Length 360;
  Best Local Similarity   29.8%;  Pred. No. 1e-10;
  Matches   97;  Conservative   56;  Mismatches  142;  Indels   31;  Gaps    7;

Qy        412 QQLLQEGIS-LGNSEADRQLLEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAA 468
              :::| : || |   : ::  |||:| |  | |:||       || :|  |:  | |  ||
Db         52 EKILMDFISNLFGRDKNQSFLEASKQGQTENVEKLLRSNKVDVNYQDAYGK--TALISAA 109

Qy        469 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 528
                    |:  |:: | ::  :|: |   | :|      :: :||||:|| :|  |     |
Db        110 DKGYRDVIGLLIESGPNLDLQDENGNTALISAAKIERGDIIDLLVKNGADLNFQDENGET 169

Qy        529 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD-GDTDIQDLLRGDAALLDAA 587
              | | |  | |  |:  | |  :: :| | |    |  || :||
Db        170 ALKFAVKVGYKNIADQLIDAGTDLNIQDENGETALICAADRAHRDIAELL---------- 219

Qy        588 KKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGL 647
                    : :  ::| :|  |:  | |  |    :  : | |:  |||:| ||| |
```

```
Db        220 ----------IKAGADLNIQDNSGK--TALVAATKIGHKGIVELLVNAGADLNLQDKNGN 267

Qy        648 IPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTL 707

Db        268 TALIYAADRGYRDIVNLLIEGGASLNIPDEAGLTALMFSAQTGRKDIVELLIKAGADINI 327

Qy        708 KNQEGQTPLDL---VSADDVSALLTA 730

Db        328 EDKNNKTAADLAAEVGFEEIVDLLTS 353


RESULT 39
ORP1_HUMAN
ID   ORP1_HUMAN     STANDARD;      PRT;   950 AA.
AC   Q9BXW6; Q9BZF5; Q9NW87;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Oxysterol binding protein-related protein 1 (OSBP-related protein 1)
DE   (ORP-1).
GN   OSBPL1A OR OSBPL1 OR ORP1 OR OSBP8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=21601154; PubMed=11735225;
RA   Jaworski C.J., Moreira E., Li A., Lee R., Rodriguez I.R.;
RT   "A family of 12 human genes containing oxysterol-binding domains.";
RL   Genomics 78:185-196(2001).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=21264638; PubMed=11279184;
RA   Xu Y., Liu Y., Ridgway N.D., McMaster C.R.;
RT   "Novel members of the human oxysterol-binding protein family bind
RT   phospholipids and regulate vesicle transport.";
RL   J. Biol. Chem. 276:18407-18414(2001).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM C).
RX   MEDLINE=21376257; PubMed=11483621;
RA   Lehto M., Laitinen S., Chinetti G., Johansson M., Ehnholm C.,
RA   Staels B., Ikonen E., Olkkonen V.M.;
RT   "The OSBP-related protein family in humans.";
RL   J. Lipid Res. 42:1203-1213(2001).
RN   [4]
RP   SEQUENCE OF 514-790 FROM N.A. (ISOFORM A).
RC   TISSUE=Embryo;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Binds phospholipids; exhibits strong binding to
CC       phosphatidic acid and weak binding to phosphatidylinositol 3-
CC       phosphate (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=B; Synonyms=OSBPL1B, OSBP8L;
CC         IsoId=Q9BXW6-1; Sequence=Displayed;
CC       Name=A; Synonyms=OSBPL1A, OSBP8S;
CC         IsoId=Q9BXW6-2; Sequence=VSP_003779;
CC       Name=C;
CC         IsoId=Q9BXW6-3; Sequence=VSP_003780;
CC   -!- SIMILARITY: Belongs to the OSBP family.
CC   -!- SIMILARITY: Contains 3 ANK repeats.
CC   -!- SIMILARITY: Contains 1 PH domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF392449; AAL40662.1; -.
DR   EMBL; AF392450; AAL40663.1; -.
DR   EMBL; AF274714; AAK15154.1; -.
DR   EMBL; AF323726; AAG53407.2; -.
DR   EMBL; AK001079; BAA91496.1; -.
DR   Genew; HGNC:16398; OSBPL1A.
DR   MIM; 606730; -.
DR   GO; GO:0005622; C:intracellular; IC.
DR   GO; GO:0005543; F:phospholipid binding; NAS.
DR   GO; GO:0008203; P:cholesterol metabolism; NAS.
DR   GO; GO:0016192; P:vesicle-mediated transport; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000648; Oxysterol_BP.
DR   InterPro; IPR001849; PH.
DR   Pfam; PF00023; ank; 3.
DR   Pfam; PF01237; Oxysterol_BP; 1.
DR   Pfam; PF00169; PH; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 3.
DR   SMART; SM00233; PH; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 3.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS01013; OSBP; 1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
KW   Lipid transport; Transport; ANK repeat; Repeat; Alternative splicing.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT       80    109       ANK 2.
FT   REPEAT      175    204       ANK 3.
FT   DOMAIN      235    334       PH.
FT   VARSPLIC      1    513       Missing (in isoform A).
FT                                /FTId=VSP_003779.
FT   VARSPLIC    625    648       FNPLLGETYELVRDDLGFRLISEQ -> L (in isoform
FT                                C).
FT                                /FTId=VSP_003780.
FT   CONFLICT    735    735       T -> A (IN REF. 4).
FT   CONFLICT    810    810       S -> P (IN REF. 1; AAL40663).
FT   CONFLICT    841    841       P -> S (IN REF. 1; AAL40662).
FT   CONFLICT    844    844       A -> G (IN REF. 1; AAL40662).
SQ   SEQUENCE   950 AA;  108469 MW;  CDDA26CA27B65F63 CRC64;

  Query Match             5.1%;  Score 291.5;  DB 1;  Length 950;
  Best Local Similarity   23.9%;  Pred. No. 7.6e-10;
  Matches  142;  Conservative   91;  Mismatches  221;  Indels  139;  Gaps   23;

Qy        423 NSEADRQLLEAAKAGDVETVKKLC-------TVQSVNCRDIEGRQSTPLHFAAGYNRVSV 475

Db          2 NTEAEQQLLHHARNGNAEEVRQLLETMARNEVIADINCK---GRSKSNL----------- 47

Qy        476 VEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAA 535

Db         48 ----------------GWTPLHLACYFGHRQVVQDLLKAGAEVNVLNDMGDTPLHRAAF 90

Qy        536 KGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG---------DAALLDA 586

Db         91 TGRKELVMLLLEYNADTTIVNGSGQTAKEVTH--AEEIRSMLEAVERTQQRKLEELLLAA 148

Qy        587 AKKGCLARVKKL---SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQD 643

Db        149 AREGKTTELTALLNRPNPPDVNCSDQLG--NTPLHCAAYRAHKQCALKLLRSGADPNLKN 206

Qy        644 KGGLIPLHNA--ASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAH 701

Db        207 KNDQKPLDLAQGAEMKHILVGNKVI-YKAL-----KRYEGPLWKSSRFFGWRLFWVVLEH 260

Qy        702 GADPTLKNQ---------EGQTPLDLVSADDVSALLTAAMPPSAL--PSCYKPQVLNGVR 750

Db        261 GVLSWYRKQPDAVHNIYRQGCKHL-------TQAVCTVKSTDSCLFFIKCF-DDTIHGFR 312
```

```
Qy        751 SP-----GATADALSSGPSSPSSLSAASSLDNLSGSFSE--LSSVVSSSGTEGASSLEKK 803
              |       :  | | :     :  :    | | |:    |  :|:        | | | :::
Db        313 VPKNSLQQSREDWLEAIEEHSAYSTHYCSQDQLTDEEEEDTVSAADLKKSLEKAQSCQQR 372

Qy        804 EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKG 863
                :|  |:|:|:: :                  |:  ||    |::: : :   |   :
Db        373 ----LDREISNFLKMIK-------------ECDMAKEMLPSFLQKVEVVSEASRETCV-- 413

Qy        864 VERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRD-------G 916
                      |  | |    |       |  : :: : : | :::    || :         |
Db        414 ---------ALTDCLNLFTKQEGVRNFKLEQEQEKNKILSEALETLATEHHELEQSLVKG 464

Qy        917 GHAGGIFNR---YNIL----------KIQKVCNKKLWERYTH---RRKEVSEE 953
                  |  :    |: |           ::: |  :   |   |   |:  :|||
Db        465 SPPASILSEDEFYDALSDSESERSLSRLEAVTARSFEEEGEHLGSRKHRMSEE 517


RESULT 40
KBF2_HUMAN
ID   KBF2_HUMAN     STANDARD;      PRT;   899 AA.
AC   Q00653; Q04860; Q9BU75; Q9H471;
DT   01-OCT-1993 (Rel. 27, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Nuclear factor NF-kappa-B p100/p49 subunits (H2TF1) (Oncogene Lyt-10)
DE   (Lyt10) [Contains: Nuclear factor NF-kappa-B p52 subunit].
GN   NFKB2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Leukemia;
RX   MEDLINE=91343004; PubMed=1876189;
RA   Schmid R.M., Perkins N.D., Duckett C.S., Andrews P.C., Nabel G.J.;
RT   "Cloning of an NF-kappa B subunit which stimulates HIV transcription
RT   in synergy with p65.";
RL   Nature 352:733-736(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92123193; PubMed=1531086;
RA   Bours V., Burd P.R., Brown K., Villalobos J., Park S., Ryseck R.P.,
RA   Bravo R., Kelly K., Siebenlist U.;
RT   "A novel mitogen-inducible gene product related to p50/p105-NF-kappa
RT   B participates in transactivation through a kappa B site.";
RL   Mol. Cell. Biol. 12:685-695(1992).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM P80HT).
RX   MEDLINE=94309910; PubMed=8036016;
RA   Thakur S., Lin H.C., Tseng W.T., Kumar S., Bravo R., Foss F.,
RA   Gelinas C., Rabson A.B.;
RT   "Rearrangement and altered expression of the NFKB-2 gene in human
RT   cutaneous T-lymphoma cells.";
RL   Oncogene 9:2335-2344(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Wilson S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE OF 1-220 FROM N.A.
RX   MEDLINE=95059001; PubMed=7969113;
RA   Liptay S., Schmid R.M., Nabel E.G., Nabel G.J.;
RT   "Transcriptional regulation of NF-kappa B2: evidence for kappa B-
RT   mediated positive and negative autoregulation.";
RL   Mol. Cell. Biol. 14:7695-7703(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 37-327.
RX   MEDLINE=98046037; PubMed=9384586;
RA   Cramer P., Larson C.J., Verdine G.L., Mueller C.W.;
RT   "Structure of the human NF-kappaB p52 homodimer-DNA complex at 2.1-A
RT   resolution.";
RL   EMBO J. 16:7078-7090(1997).
CC   -!- FUNCTION: P100 is the precursor of the p52 subunit of the nuclear
CC       factor NF-kappa-B, which binds to the kappa-B consensus sequence
CC       5'-GGRNNYYCC-3', located in the enhancer region of genes involved
CC       in immune response and acute phase reactions. The precursor
CC       protein itself does not bind to DNA. Isoform p49 is a subunit of
CC       the NF-kappa-B protein complex, which stimulates the HIV enhancer
CC       in synergy with p65.
CC   -!- SUBUNIT: Active NF-kappa-B is a heterodimer of an about 52 kDa
CC       DNA-binding subunit and the weak DNA-binding subunit p65. Two
CC       heterodimers might form a labile tetramer.
CC   -!- SUBCELLULAR LOCATION: Nuclear, but also found in the cytoplasm in
CC       an inactive form complexed to an inhibitor (I-kappa-B).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=p100;
CC         IsoId=Q00653-1; Sequence=Displayed;
CC       Name=p49;
CC         IsoId=Q00653-2; Sequence=VSP_005585, VSP_005586;
CC   -!- DOMAIN: The C-terminus of p100 might be involved in cytoplasmic
CC       retention, inhibition of DNA-binding by p52 homodimers, and/or
CC       transcription activation (By similarity).
CC   -!- DISEASE: In cutaneous T-cell leukemia (CTCL) cell line HUT 78, a
CC       rearranged NFKB2 gene is found which encodes for a truncated 80
CC       kDa protein (p80HT).
CC   -!- SIMILARITY: Belongs to the Rel/Dorsal family.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- SIMILARITY: Contains 1 death domain.
CC   -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC       WWW="http://www.infobiogen.fr/services/chromcancer/Genes/NFKB2ID362.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61498; CAA43715.1; -.
DR   EMBL; X61499; CAA43716.1; ALT_SEQ.
DR   EMBL; S76638; AAB21124.1; -.
DR   EMBL; U09609; AAA21462.1; -.
DR   EMBL; AL121928; CAC08399.1; -.
DR   EMBL; BC002844; AAH02844.1; -.
DR   EMBL; U20816; AAA68171.1; -.
DR   PIR; A42024; A42024.
DR   PIR; S17233; S17233.
DR   PDB; 1A3Q; 27-MAY-98.
DR   TRANSFAC; T01927; -.
```

```
DR   TRANSFAC; T01930; -.
DR   Genew; HGNC:7795; NFKB2.
DR   MIM; 164012; -.
DR   GO; GO:0003713; F:transcription co-activator activity; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR002909; IPT_TIG.
DR   InterPro; IPR000451; NF_Rel_dor.
DR   InterPro; IPR008967; P53-like.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00554; RHD; 1.
DR   Pfam; PF01833; TIG; 1.
DR   PRINTS; PR00057; NFKBTNSCPFCT.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00429; IPT; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR   PROSITE; PS01204; REL_1; 1.
DR   PROSITE; PS50254; REL_2; 1.
KW   DNA-binding; Transcription regulation; Activator; Nuclear protein;
KW   Repeat; ANK repeat; Phosphorylation; Alternative splicing;
KW   Disease mutation; Proto-oncogene; 3D-structure.
FT   DOMAIN       38    343       REL-LIKE (RHD).
FT   REPEAT      487    519       ANK 1.
FT   REPEAT      526    555       ANK 2.
FT   REPEAT      559    591       ANK 3.
FT   REPEAT      599    628       ANK 4.
FT   REPEAT      633    663       ANK 5.
FT   REPEAT      667    696       ANK 6.
FT   REPEAT      729    758       ANK 7.
FT   DOMAIN      764    851       DEATH.
FT   DOMAIN      337    341       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      350    400       GLY-RICH.
FT   VARSPLIC    374    447       GSLGFFPSSLAYSPYQSGAGPMGCYPGGGGGAQMAATVPSR
FT                                DSGEEAAEPSAPSRTPQCEPQAPEMLQRAREYN -> EGVL
FT                                MEGGVKVREAVEEKNLGEAGRGLHACNPAFGRPRQADYLRS
FT                                GVQDQLGQQRETSSLLKIQTLAGHGGRRL (in isoform
FT                                p49).
FT                                /FTId=VSP_005585.
FT   VARSPLIC    448    899       Missing (in isoform p49).
FT                                /FTId=VSP_005586.
FT   VARIANT     667    669       AGN -> SAS (IN TRUNCATED FORM P80HT).
FT                                /FTId=VAR_006909.
FT   VARIANT     670    899       MISSING (IN TRUNCATED FORM P80HT).
FT                                /FTId=VAR_006910.
FT   CONFLICT    144    144       K -> E (IN REF. 2).
FT   CONFLICT    213    213       I -> T (IN REF. 1, 3 AND 6).
FT   CONFLICT    396    396       G -> R (IN REF. 1 AND 3).
FT   CONFLICT    433    434       EP -> DA (IN REF. 2).
FT   CONFLICT    455    470       QRSARALLDYGVTADA -> HAAPSPTRLLRHRGR (IN
FT                                REF. 1 AND 3).
FT   CONFLICT    470    470       A -> R (IN REF. 2).
FT   CONFLICT    741    741       K -> L (IN REF. 1).
FT   CONFLICT    859    859       T -> TA (IN REF. 4).
FT   CONFLICT    859    859       T -> TT (IN REF. 2).
FT   CONFLICT    875    875       E -> K (IN REF. 2).
FT   CONFLICT    890    890       C -> S (IN REF. 1).
FT   CONFLICT    899    899       H -> TDLLPAPSPLPGPPVQRPHLFQILFNTPHPPLSWDK
FT                                (IN REF. 1).
FT   STRAND       39     44
FT   STRAND       46     46
FT   STRAND       51     51
FT   STRAND       54     55
FT   TURN         56     58
FT   STRAND       67     67
FT   TURN         72     73
FT   STRAND       79     83
FT   TURN         84     84
```

```
  Query Match           5.0%;  Score 289;  DB 1;  Length 899;
  Best Local Similarity   26.5%;  Pred. No. 9.9e-10;
  Matches  131;  Conservative   60;  Mismatches  181;  Indels  122;  Gaps   19;

Qy        406 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 465
              :   : : ||  |::      || : | | :       : | | | |         ||||
Db        453 LAQRSARALLDYGVT-----ADARALLAGQ-------RHLLTAQDEN-------GDTPLH 493

Qy        466 FAAGYNRVSVVE---YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 522
              |  : : ||:|   |:: |  |:                                |||:
Db        494 LAIIHGQTSVIEQIVYVIHHAQDL----------------------------GVVNLT 523

Qy        523 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAA 582
              :    |||| |  |:  :   ||: ||||   :| |:: : |          :|||   |
Db        524 NHLHQTPLHLAVITGQTSVVSFLLRVGADPALLDRHGDSAMHLALRAGAGAPELLR---A 580

Qy        583 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642
              || :           : |  ::  | :| :  |:|||    : |  : |:  ||:| |
Db        581 LLQSGAP---------AVPQLLHMPDFEGLY--PVHLAVRARSPECLDLLVDSGAEVEAT 629

Qy        643 DK-GGLIPLHNAASYGHVD-VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 700
              :: ||   || |      :  |  |: |  | |||       |||| ||  |   |  |||
Db        630 ERQGGRTALHLATEMEELGLVTHLVTKLRANVNARTFAGNTPLHLAAGLGYPTLTRLLLK 689

Qy        701 HGADPTLKNQE----------------------------GQTPLDLVSADDVSALLTAA 731
              |||   :|:|                            | |||||  :  |  ||  |
Db        690 AGADIHAENEEPLCPLPSPPTSDSDSDSEGPEKDTRSSFRGHTPLDLTCSTKVKTLLLNA 749

Qy        732 M-----PPSALPSCYKPQVLNGVRSPGATA----DALSSGPSSPSSLSAASSLDNLSGSF 782
                  ||   ||   | :    | | ||    : |  || :  | :  :    |
Db        750 AQNTMEPPLTPPSPAGPGL-----SLGDTALQNLEQLLDGPEAQGSWAELAERLGLRSLV 804

Qy        783 SELSSVVSSSGTEGASSLEKKEVPGVDFS-ITQFVRNLGLEHLMDIFEREQITLDVLVEM 841
                  | ||    | |   |: | | : : : : ::|||  : :  |  | |   ::
Db        805 DTYRQTTSPSG----SLLRSYELAGGDLAGLLEALSDMGLEEGVRLL-RGPETRD---KL 856

Qy        842 GHKELKEIGINAYG 855
                |:||   :|||
Db        857 PSTEVKE--DSAYG 868
```

```
RESULT 41
FEM1_CAEEL
ID   FEM1_CAEEL     STANDARD;      PRT;   656 AA.
AC   P17221;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Sex-determining protein fem-1.
GN   FEM-1 OR ISX-1 OR F35D6.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=90199879; PubMed=2317869;
RA   Spence A.M., Coulson A., Hodgkin J.;
RT   "The product of fem-1, a nematode sex-determining gene, contains a
RT   found in cell cycle control proteins and receptors for
RT   cell-cell interactions.";
RL   Cell 60:981-990(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Sex-determining protein. It is essential for the
CC       adoption of the male sexual fate in all tissues in C.elegans.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Contains 7 ANK repeats.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J03172; AAA28055.1; -.
DR   EMBL; U52001; AAA96093.1; -.
DR   PIR; A34793; A34793.
DR   HSSP; Q00420; 1AWC.
DR   WormPep; F35D6.1; CE07175.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 8.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 5.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 2.
KW   Developmental protein; Phosphorylation; ANK repeat; Repeat.
FT   REPEAT       47     76       ANK 1.
FT   REPEAT       88    118       ANK 2.
FT   REPEAT      122    151       ANK 3.
FT   REPEAT      155    184       ANK 4.
FT   REPEAT      188    217       ANK 5.
FT   REPEAT      220    250       ANK 6.
FT   REPEAT      557    588       ANK 7.
SQ   SEQUENCE   656 AA;  74286 MW;  4F5FC3FA3EA950D3 CRC64;

  Query Match             5.0%;  Score 287.5;  DB 1;  Length 656;
  Best Local Similarity   22.8%;  Pred. No. 7.6e-10;
  Matches  170;  Conservative   87;  Mismatches  256;  Indels  233;  Gaps   28;

Qy          3 GRKDVVEYLLQNGASVQAR--------DDGGLIPLHNACSFGHAEVVNLLL-RHGADPNA 53
              |  :||||||: ||     |         :  |  ||  | : || |:| ||: :  || |
Db         59 GHANVVEYLLEIGADPSVRGVVEFDNENIQGTPPLWAASAAGHIEIVKLLIEKANADVNQ 118

Qy         54 RDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKD 113
               |  |||  |   | :|:   ||: ||:| | |  | | | :|          |:
Db        119 ATNTRSTPLRGACYDGHLDIVKYLLEKGADPHIPNRHGHTCLMIA----------SYR-- 166

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHA 173
                                                  |:| ||: ||: | ||:
Db        167 ----------------------------------------NKVGIVEELLKTGIDVNK 184

Qy        174 KDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSY 233
              | :    ||:|   |: || ::|:|||: |   |:    ||  ||   ::| ::|
Db        185 KTERGNTALHDAAESGNVEVVKILLKHGS-VLMKDIQGVDPLMGAALSGFLDVLNVL--- 240

Qy        234 GADPTLLNCHNKSAIDLAPTPQLKERL----AYEFKGHSLLQAAREADVTRIKKHLSLEM 289
               ||      | : |:|  :  | :::     |     | |: |  ::    ||  |: :   ||
Db        241 -ADQMPSGIHKRDALKLLGSTLLDKKMDAMSAMNCWKQS-MEVPLHADDLRLVR----EM 294

Qy        290 VNFKHPQTHETALHCAAASPYPKRKQ-ICELLLRKGANINEKTKEFLTPLHVASEKAHND 348
              |  ||           | :  | | :: |  | ||  :  : |          || |
Db        295 ETFFEPQ---------EVYEYQREAQNIAQVELLDG-NIEAQRMQALVIRERILGGAHTD 344

Qy        349 VVEVVVKHEAKVNALDNLGQTSLHRAAYC--GHLQTCRLLLSYGCDPNIISLQGFTALQM 406
              |  :                  | ||  | :  |  |  :  :   :  : |  |
Db        345 VHYYL----------------RFRGAVYCDMGQMNRCYDLWKHALE---LQQKHFAPLYY 385

Qy        407 GNENVQQLLQEGISL------GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 460
              |     |  |  |:        |  |:| |
Db        386 GTITTLQSFHETFSMSLNDFVNNHHANRNLRV--------------------------- 417

Qy        461 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 520
                        | | |:|:                |  :  |  | :
Db        418 ----------RSSWVKYVF----------NGVCLELERAAA----------------- 438

Qy        521 VADLWKFTPLHEAAAKGKYEIC-KLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 579
                  |   || |        | | |   ||  : ::  :     : ||     :: : |
Db        439 ----WTGAPLLEDT-----ECCGKERCQHATEESEYKKLVYVAVHLV-----NVLERLSL 484

Qy        580 DAALLDAA--KKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGA 637
                :| :|  :  :|    | |::|      : | :          ||:   :: ||
Db        485 PSAHIDDSDEEKEPKADVRRLM----IVCHE--------LHIPLLHHTLEER-------- 524

Qy        638 DVNAQDKGGLIPLHNAASYGHVDVAAL--LIKYNACVNATDKWAFTPLH--EAAQKGRTQ 693
                        ||  |:|  |    | |  |:: :  ||||||   ||:|    |:: |
Db        525 ----------IPDSNSAELGLPKAAVLEQLLELDLDVNATDKNNDTPMHILLRAREFRKS 574

Qy        694 LCALLLAHGADPTLKNQEGQTPLDLV 719
              |   ||  |      :|: |   |:::
Db        575 LVRALLVRGTWLFARNRHGDVVLNVM 600
```

```
RESULT 42
RN5A_HUMAN
ID   RN5A_HUMAN     STANDARD;      PRT;   741 AA.
AC   Q05823;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   2-5A-dependent ribonuclease (EC 3.1.26.-) (2-5A-dependent RNase)
DE   (Ribonuclease L) (RNase L) (Ribonuclease 4).
GN   RNASEL OR RNS4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-240 AND LYS-274.
RC   TISSUE=Kidney;
RX   MEDLINE=93201598; PubMed=7680958;
RA   Zhou A., Hassel B.A., Silverman R.H.;
RT   "Expression cloning of 2-5A-dependent RNAase: a uniquely regulated
RT   mediator of interferon action.";
RL   Cell 72:753-765(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=20515260; PubMed=11063255;
RA   Zhou A., Nie H., Silverman R.H.;
RT   "Analysis and origins of the human and mouse RNase L genes: mediators
RT   of interferon action.";
RL   Mamm. Genome 11:989-992(2000).
RN   [3]
RP   CHARACTERIZATION OF RNASEL ACTIVITY.
RX   MEDLINE=94245737; PubMed=7514601;
RA   Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,
RA   Silverman R.H.;
RT   "Intrinsic molecular activities of the interferon-induced 2-5A-
RT   dependent RNase.";
RL   J. Biol. Chem. 269:14153-14158(1994).
RN   [4]
RP   FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21611192; PubMed=11585831;
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RT   "The 2-5A/RNase L/RNase L inhibitor (RNI) pathway regulates
RT   mitochondrial mRNAs stability in interferon alpha-treated H9 cells.";
RL   J. Biol. Chem. 276:48473-48482(2001).
RN   [5]
RP   ERRATUM.
RA   Le Roy F., Bisbal C., Silhol M., Martinand C., Lebleu B.,
RA   Salehzada T.;
RL   J. Biol. Chem. 277:13354-13354(2002).
RN   [6]
RP   REVIEW.
RX   MEDLINE=99073516; PubMed=9856285;
RA   Castelli J., Wood K.A., Youle R.J.;
RT   "The 2-5A system in viral infection and apoptosis.";
RL   Biomed. Pharmacother. 52:386-390(1998).
RN   [7]
RP   MUTAGENESIS OF LYS-392.
```

```
RX   MEDLINE=99081839; PubMed=9862963;
RA   Dong B., Silverman R.H.;
RT   "Alternative function of a protein kinase homology domain in 2',
RT   5'-oligoadenylate dependent RNase L.";
RL   Nucleic Acids Res. 27:439-445(1999).
RN   [8]
RP   MUTAGENESIS OF HIS-583; PRO-584; TRP-632; ASP-661; ARG-667 AND
RP   HIS-672.
RX   MEDLINE=21230542; PubMed=11333017;
RA   Dong B., Niwa M., Walter P., Silverman R.H.;
RT   "Basis for regulated RNA cleavage by functional analysis of RNase L
RT   and Irelp.";
RL   RNA 7:361-373(2001).
RN   [9]
RP   VARIANTS SER-59; PHE-406; GLN-462 AND GLU-541.
RX   MEDLINE=21947464; PubMed=11941539;
RA   Roekman A., Ikonen T., Seppaelae E.H., Nupponen N., Autio V.,
RA   Mononen N., Bailey-Wilson J., Trent J., Carpten J., Matikainen M.P.,
RA   Koivisto P.A., Tammela T.L.J., Kallioniemi O.-P., Schleutker J.;
RT   "Germline alterations of the RNASEL gene, a candidate HPC1 gene at
RT   1q25, in patients and families with prostate cancer.";
RL   Am. J. Hum. Genet. 70:1299-1304(2002).
RN   [10]
RP   VARIANTS SER-59; GLN-462 AND GLU-541.
RX   MEDLINE=21676832; PubMed=11799394;
RA   Carpten J., Nupponen N., Isaacs S., Sood R., Robbins C., Xu J.,
RA   Faruque M., Moses T., Ewing C., Gillanders E., Hu P., Bujnovszky P.,
RA   Makalowska I., Baffoe-Bonnie A., Faith D., Smith J., Stephan D.,
RA   Wiley K., Brownstein M., Gildea D., Kelly B., Jenkins R.,
RA   Hostetter G., Matikainen M., Schleutker J., Klinger K., Connors T.,
RA   Xiang Y., Wang Z., De Marzo A., Papadopoulos N., Kallioniemi O.-P.,
RA   Burk R., Meyers D., Groenberg H., Meltzer P., Silverman R.,
RA   Bailey-Wilson J., Walsh P., Isaacs W., Trent J.;
RT   "Germline mutations in the ribonuclease L gene in families showing
RT   linkage with HPC1.";
RL   Nat. Genet. 30:181-184(2002).
RN   [11]
RP   CHARACTERIZATION OF VARIANTS GLN-462 AND GLU-541.
RX   MEDLINE=22345000; PubMed=12415269;
RA   Casey G., Neville P.J., Plummer S.J., Xiang Y., Krumroy L.M.,
RA   Klein E.A., Catalona W.J., Nupponen N., Carpten J.D., Trent J.M.,
RA   Silverman R.H., Witte J.S.;
RT   "RNASEL Arg462Gln variant is implicated in up to 13% of prostate
RT   cancer cases.";
RL   Nat. Genet. 32:581-583(2002).
CC   -!- FUNCTION: Endoribonuclease, mediator of interferon action, which
CC       play a role in mediating resistance to virus infection and
CC       apoptosis. Might play a central role in the regulation of mRNA
CC       turnover.
CC   -!- CATALYTIC ACTIVITY: Cleaves 3' of UpNp dimers, with preference for
CC       UU and UA sequences, to sets of discrete products ranging from
CC       between 4 and 22 nucleotides in length.
CC   -!- COFACTOR: Optimal RNA cleavage rates requires the presence of
CC       either manganese or magnesium and ATP.
CC   -!- ENZYME REGULATION: After binding to 2-5A (5'-phosphorylated 2',5'-
CC       linked oligoadenylates) the homodimerization and subsequent
CC       activation occurs. Inhibited by RNase L inhibitor.
CC   -!- SUBUNIT: Monomer (inactive form) or homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.
CC   -!- TISSUE SPECIFICITY: Highly expressed in spleen and thymus followed
CC       by prostate, testis, uterus, small intestine, colon and peripheral
CC       blood leukocytes.
CC   -!- INDUCTION: By interferons.
CC   -!- DOMAIN: The kinase domain allows the homodimerization.
CC   -!- DOMAIN: The nine ankyrin repeats also called 2-5A sensor
CC       constitute the 2-5A binding domain.
CC   -!- DISEASE: Defects in RNASEL may be the cause of susceptibility to
CC       hereditary prostate cancer 1 (HPC1) [MIM:601518].
CC   -!- SIMILARITY: Contains 9 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10381; AAA18032.1; -.
DR   HSSP; P80144; 2MYO.
DR   Genew; HGNC:10050; RNASEL.
DR   MIM; 180435; -.
DR   MIM; 601518; -.
DR   GO; GO:0004521; F:endoribonuclease activity; NAS.
DR   GO; GO:0003723; F:RNA binding; IDA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR006567; PUG.
DR   Pfam; PF00023; ank; 9.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00248; ANK; 8.
DR   SMART; SM00580; PUG; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 6.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Hydrolase; Nuclease; RNA-binding; Endonuclease; Zinc-finger; Repeat;
KW   ANK repeat; Polymorphism.
FT   REPEAT      24     53      ANK 1.
FT   REPEAT      58     87      ANK 2.
FT   REPEAT      91    120      ANK 3.
FT   REPEAT     124    153      ANK 4.
FT   REPEAT     167    197      ANK 5.
FT   REPEAT     201    234      ANK 6.
FT   REPEAT     238    268      ANK 7.
FT   REPEAT     272    301      ANK 8.
FT   REPEAT     303    329      ANK 9.
FT   DOMAIN     365    586      PROTEIN KINASE LIKE.
FT   DOMAIN     587    741      RIBONUCLEASE.
FT   DOMAIN     229    242      2-5A BINDING (P-LOOP) 1.
FT   DOMAIN     253    275      2-5A BINDING (P-LOOP) 2.
FT   ZN_FING    395    444      C6-TYPE (ATYPICAL).
FT   VARIANT     59     59      G -> S.
FT                              /FTId=VAR_013509.
FT   VARIANT    406    406      S -> F.
FT                              /FTId=VAR_013510.
FT   VARIANT    462    462      R -> Q (risk factor for prostate cancer;
FT                              reduced enzymatic activity;
FT                              dbSNP:486907).
FT                              /FTId=VAR_012056.
FT   VARIANT    541    541      D -> E (no change in enzymatic activity;
FT                              dbSNP:627928).
FT                              /FTId=VAR_012057.
FT   MUTAGEN    240    240      K->N: REDUCED 2-5A BINDING ACTIVITY;
FT                              ALMOST COMPLETE LOSS OF 2-5A BINDING
FT                              ACTIVITY; WHEN ASSOCIATED WITH N-274.
FT   MUTAGEN    274    274      K->N: REDUCED 2-5A BINDING ACTIVITY;
FT                              ALMOST COMPLETE LOSS OF 2-5A BINDING
FT                              ACTIVITY; WHEN ASSOCIATED WITH N-240.
FT   MUTAGEN    392    392      K->R: COMPLETE LOSS OF ENZYMATIC ACTIVITY
FT                              AND ENZYME DIMERIZATION. NO CHANGE IN
FT                              BINDING TO 2-5A AND RNA.
FT   MUTAGEN    583    583      H->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN    584    584      P->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN    632    632      W->A: NO CHANGE IN ENZYMATIC ACTIVITY.
FT   MUTAGEN    661    661      D->A: COMPLETE LOSS OF ENZYMATIC
FT                              ACTIVITY.
FT   MUTAGEN    667    667      R->A: COMPLETE LOSS OF ENZYMATIC
FT                              ACTIVITY. NO CHANGE IN 2-5A BINDING AND
FT                              ENZYME DIMERIZATION.
FT   MUTAGEN    672    672      H->A: COMPLETE LOSS OF ENZYMATIC
FT                              ACTIVITY. NO CHANGE IN 2-5A BINDING

  Query Match             5.0%;  Score 286.5;  DB 1;  Length 741;
  Best Local Similarity  23.4%;  Pred. No. 1.1e-09;
```

```
  Matches  148;  Conservative   84;  Mismatches  210;  Indels  191;  Gaps   24;

Qy          6 DVVEYLLQNGASVQ-ARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHE 64
              |:|: ||: ||:|    ::|| |||||      ::| |||||||||  |      ||
Db         39 DLVQQLLEGGANVNFQEEEGGWTPLHNAVQMSREDIVELLLRHGADPVLRKKNGATPFIL 98

Qy         65 AAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNE 124
              ||| | : :  : |  ||:    :  | ||                 :|
Db         99 AAIAGSVKLLKLFLSKGADVNECDFYGFTA--------------------FME------- 131

Qy        125 EKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK---------- 174
                                      || | :|| :: | : ||:|: :
Db        132 --------------------------AAVYGKVKALKFLYKRGANVNLRRKTKEDQERL 164

Qy        175 DKGDLVPLHNACSYGHYEVTELLV-KHGACVNAMD-LWQFTPLHEAASKNRVEV---CSL 229
               ||    | :|   || || ::|: : || ||| | : :   :|   | :  :|     |
Db        165 RKGGATALMDAAEKGHVEVLKILLDEMGADVNACDNMGRNALIHALLSSDDSDVEAITHL 224

Qy        230 LLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEM 289
              || :|||  :     |       || :   || | |   |:|   |      ::|: :
Db        225 LLDHGADVNVRGERGK-------TPLI---LAVEKKHLGLVQRLLE------QEHIEIND 268

Qy        290 VNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDV 349
               :  :|||  |      | |:| ||| ::||     : :   :  |     : :
Db        269 TD----SDGKTALLLAVEL---KLKKIAELLCKRGA-----STDCGDLVMTARRNYDHSL 316

Qy        350 VEVVVKHEAKVN----ALDNLGQTS--------LHR------------------------ 373
              |:|:: | || :    | |   |:|         |||
Db        317 VKVLLSHGAKEDFHPPAEDWKPQSSHWGAALKDLHRIYRPMIGKLKFFIDEKYKIADTSE 376

Qy        374 -AAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 432
               | |  :   : :   |: :  : :  : :|   ||     ||        ::   ||: |
Db        377 GGIYLGFYEKQEVAVKTFCEGSPRAQREVSCLQSSREN-----SHLVTFYGSESHR---- 427

Qy        433 AAKAGDVETVKKLCTVQSVNCRDI---EGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAK 489
               | :     ||       | |:   |  ::    ||   | :| :   :|
Db        428 ----GHLFVCVTLCEQTLEACLDVHRGEDVENEEDEFAR--NVLSSIFKAVQE------- 474

Qy        490 DKGGLVPLHNACSYGHYEV--AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQ 547
                  || :| | | ::    :|:      ::||  |                    ::
Db        475 -------LHLSCGYTHQDLQPQNILIDSKKAAHLADFDK-----------------SIK 509

Qy        548 HGADPTKKNRD----GNTPLDLVKDGDTDIQDL 576
               || :  ||    |  | :|| |    :||
Db        510 WAGDPQEVKRDLEDLGRLVLYVVKKGSISFEDL 542


RESULT 43
V218_FOWPV
ID   V218_FOWPV     STANDARD;      PRT;   461 AA.
AC   Q9J517;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative ankyrin-repeat protein FPV218.
GN   FPV218.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
CC   -!- SIMILARITY: Contains 12 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF198100; AAF44562.1; -.
DR   HSSP; P80144; 2MYO.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 9.
DR   SMART; SM00248; ANK; 9.
DR   PROSITE; PS50088; ANK_REPEAT; 3.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; Repeat; ANK repeat.
FT   REPEAT        1     28       ANK 1.
FT   REPEAT       31     61       ANK 2.
FT   REPEAT       65     94       ANK 3.
FT   REPEAT       96    116       ANK 4.
FT   REPEAT      120    149       ANK 5.
FT   REPEAT      153    182       ANK 6.
FT   REPEAT      186    213       ANK 7.
FT   REPEAT      217    248       ANK 8.
FT   REPEAT      250    277       ANK 9.
FT   REPEAT      281    312       ANK 10.
FT   REPEAT      358    385       ANK 11.
FT   REPEAT      431    460       ANK 12.
SQ   SEQUENCE   461 AA;  52636 MW;  E5892ABD916AB807 CRC64;

  Query Match             4.9%;  Score 285;  DB 1;  Length 461;
  Best Local Similarity   28.6%;  Pred. No. 6.4e-10;
  Matches   98;  Conservative   54;  Mismatches  141;  Indels   50;  Gaps   11;

Qy        411 VQQLLQEGISLGNSEAD--RQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAA 468
              |::|||:||:   ::      | |:::    ||: |  |       |  :|     :||  |
Db         16 VERLLREGVHPDSTIKGFYRPLVKSILLRDVDLVSILLQ-NGANPNNINDETVSPLAIAI 74

Qy        469 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 528
               |  ::|  || : ||        | ||      |  : : ::|| ||  ||| |    :
Db         75 KVNSPTIVSLLLDYNADT------SLFPL-----YVSFPIIKVLVYHGIDVNVIDRESRS 123

Qy        529 PLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 588
               || ||     :   |: |||:   ::  | :||                    :|
Db        124 FLHYAAKNDDVDTVISLILHGANVNVQDSKGLSPLHHA------------------VSK 164

Qy        589 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 648
              | |      | :   || ||: ||   ||||  | |  |: : |: :|: :: :|  |
Db        165 KTTLTAKILLENGARVNIRDSLGR--LPLHL--GANTYEMVKLLIDYGSPIDIKDVNGST 220

Qy        649 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 708
              ||| |     :|   ||:  :  :|| |    :|||        |::   ||  ||| |:|
Db        221 PLHYAIWKSSLDTIRLLVNVST-INALDNNCNSPLHYII-LSETEILVELLLRGADITIK 278

Qy        709 NQEGQTPLDLV------SADDVSALLTAA-----MPPSALPSC 740
              :  | ||||::        |:: |::: |      :  |  |  |
Db        279 DICGNTPLDILCKLRIKKLDNIKAIISNAFLMREVVPDLLKLC 321


RESULT 44
HOS4_YEAST
ID   HOS4_YEAST     STANDARD;      PRT;  1083 AA.
AC   P40480;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein HOS4.
GN   HOS4 OR YIL112W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313266; PubMed=9169870;
RA   Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA   Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
```

```
RA   Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
RA   Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA   Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL   Nature 387:84-87(1997).
RN   [2]
RP   IDENTIFICATION IN A COMPLEX WITH HOS2; HST1; SNT1; SIF2; CPR1 AND
RP   SET3.
RX   MEDLINE=21567937; PubMed=11711434;
RA   Pijnappel W.W.M.P., Schaft D., Roguev A., Shevchenko A., Tekotte H.,
RA   Wilm M., Rigaut G., Seraphin B., Aasland R., Stewart A.F.;
RT   "The S. cerevisiae SET3 complex includes two histone deacetylases,
RT   Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation
RT   gene program.";
RL   Genes Dev. 15:2991-3004(2001).
CC   -!- FUNCTION: Unknown. Component of the Set3C complex, which is
CC       required to repress early/middle sporulation genes during meiosis.
CC   -!- SUBUNIT: Identified in the Set3C complex with HOS2, HST1, SNT1,
CC       SIF2, CPR1 and SET3.
CC   -!- SIMILARITY: Contains 5 ANK repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z38125; CAA86268.1; -.
DR   PIR; S48460; S48460.
DR   HSSP; P80144; 2MYO.
DR   GermOnline; 139647; -.
DR   SGD; S0001374; HOS4.
DR   GO; GO:0000118; C:histone deacetylase complex; IPI.
DR   GO; GO:0017136; F:NAD-dependent histone deacetylase activity; IDA.
DR   GO; GO:0045129; F:NAD-independent histone deacetylase activity; IDA.
DR   GO; GO:0016575; P:histone deacetylation; IDA.
DR   GO; GO:0045835; P:negative regulation of meiosis; IPI.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 6.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 4.
DR   PROSITE; PS50088; ANK_REPEAT; 4.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   REPEAT      329    359       ANK 1.
FT   REPEAT      363    392       ANK 2.
FT   REPEAT      398    427       ANK 3.
FT   REPEAT      532    561       ANK 4.
FT   REPEAT      593    622       ANK 5.
SQ   SEQUENCE   1083 AA;  123556 MW;  D73BADE712419555D CRC64;

  Query Match             4.9%;  Score 285;  DB 1;  Length 1083;
  Best Local Similarity   20.3%;  Pred. No. 2.3e-09;
  Matches  192;  Conservative  135;  Mismatches  365;  Indels  254;  Gaps   33;

Qy        211 QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL 270
              | | :|:|   |              ||   | : |||                   |
Db         71 QTTSVHQATQNN----------------LNA--KESEDLA--------------HKND 96

Qy        271 QAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASP--------------------- 309
               :: | :|    :   :   | |  |     :  :::||
Db         97 ASSHEGEVNGDSRPDDVPETNEKISQAIRAKISSSSSSPNVRNVDIQNHQPFSRDQLRAM 156

Qy        310 --YPKRKQICELLLRKGANINEK--------TKEFLTPLHVASE----KAHNDVVEVVVK 355
                 |||| : : :  :|    |:         |    | :|  :      :   |  :|
Db        157 LKEPKRKTVDDFIEEEGLGAVEEEDLSDEVLEKNTTEPENVEKDIEYSDSDKDTDDVGSD 216

Qy        356 HEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLL 415
                   |:   ||:  | |                    |   |:    ||:  :|
Db        217 DPTAPNSPIKLGRRKLVRGD--------------------QLDATTSSMFNNESDSELS 255

Qy        416 ----QEGISLGNS----------EADRQL-------LEAAKAGDVE----------TVK 443
                  : |:| :|          | :   |       :  | | |              |
Db        256 DIDDSKNIALSSSLFRGGSSPVKETNNNLSNMNSSPAQNPKRGSVSRSNDSNKSSHIAVS 315

Qy        444 KLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGA-DVHAKDKGGLVPLHNACS 502
              |    :   ||  ||  | |  |    :  ||: ::: |  |:: :|  |    || |
Db        316 KRPKQKKGIYRDSGGR--TRLQIACDKGKYDVVKKMIEEGGYDINDQDNAGNTALHEAAL 373

Qy        503 YGHYEVAELLVKHGAVVNVADLWKF--TPLHEAAAKGKYEICKLLLQHGADPTKKNRDGN 560
              || |: |||:::|| ||:  :  |  ||| :|:|  |  :: | ||::||||| :|  |
Db        374 QGHIEIVELLIENGADVNIKSIEMFGDTPLIDASANGHLDVVKYLLKNGADPTIRNAKGL 433

Qy        561 TPLDLVKD----GDTDIQDLLRGDAALLDAAKKGCLARV----KKLSSPDNVNCRD---- 608
              |  : | |      | : | :||      |  | |      |      |  : :| :  |
Db        434 TAFESVDDESEFDDEEDQKILREIKKRLSIAAKKWTNRAGIHNDKSKNGNNAHTIDQPPF 493

Qy        609 ---TQGRHSTPLHLAAGYNNL-EVAEYLLQHGADVNAQ---------------------- 642
                 |: ::      :  :|: | |      :  || ::
Db        494 DNTTKAKNEKAADSPSMASNIDEKAPEEEFYWTDVTSRAGKEKLFKASKEGHLPYVGTYV 553

Qy        643 DKGGLIPLHN---AASYGHVDVAALLIKYNACVNATDKWAFT-PLHEAAQKGRTQLCALL 698
              : || | | :   :   || |: :: : :   || | :   |  | |  :|      ||
Db        554 ENGGKIDLRSFFESVKCGHEDITSIFLAFGFPVNQTSRDNKTSALMVAVGRGHLGTVKLL 613

Qy        699 LAHGADPTLKNQEGQTPL--------DLVSADDVSALLTAAMPPSALPS----------- 739
              |  ||||| ::::|:| |          : ::::: :  |      |
Db        614 LEAGADPTKRDKKGRTALYYAKNSIMGITNSEEIQLIENAINNYLKKHSEDNNDDDDDDD 673

Qy        740 ----CYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 795
                  || :          |   |: |          |:     |:    |  |: ::
Db        674 NNNETYKHEKKREKTQSPILASRRSATPRIEDEEDDTRMLNLADDDFNNDRDVKESTTSD 733

Qy        796 GASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIG-INAY 854
                |:  |               |:| :: :|  :|:   |         ::|| |  ::
Db        734 SRKRLDDNE------------NVGTQYSLDWKKRKTNAL-----QDEEKLKSISPLSME 775

Qy        855 GHRHKLIKGVE--------------RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              |  |  | ||                ||   ::      |          :|  |: |:|  :
Db        776 PHSPKKAKSVEISKIHEETAAEREARLKEEEEYRKKRLEKKRKKEQELLQKLAEDEK--K 833

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              :||: : | |              : :::|   :|   |    | ||: | ::  |
Db        834 RIEEQEKQKVLE--------------MERLEKATLEK--ARKMEREKEMEEISYRRAVR 876

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYV 1006
                :   |   :|      |:::  |    |    || :   |::::|
Db        877 DLYPLGLKIIN------FNDKLDY--KRFLPLYYFVD--EKNDKFV 912


RESULT 45
BAT8_HUMAN
ID   BAT8_HUMAN     STANDARD;      PRT;  1210 AA.
AC   Q96KQ7; Q14349; Q96MH5; Q96QD0; Q9UQL8; Q9Y331;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Histone-lysine N-methyltransferase, H3 lysine-9 specific 3
DE   (EC 2.1.1.43) (Histone H3-K9 methyltransferase 3) (H3-K9-HMTase 3)
DE   (HLA-B associated transcript 8) (G9a) (NG36).
GN   BAT8 OR G9A OR NG36.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rowen L., Qin S., Madan A., Dickhoff R., Dors M., Madan A., Hicks P.,
RA   Loretz C., Ratcliffe A., Abbasi N., Shaffer T., Hood L.;
RT   "Sequence of the human major histocompatibility complex class III
RT   region.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RA   Hirakawa M., Yamaguchi H., Imai K., Shimada J., Shiina S., Tamiya G.,
RA   Oka A., Inoko H.;
RT   "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING (ISOFORM 2), AND
RP   TISSUE SPECIFICITY.
RX   MEDLINE=21564388; PubMed=11707778;
RA   Brown S.E., Campbell R.D., Sanderson C.M.;
RT   "Novel NG36/G9a gene products encoded within the human and mouse MHC
RT   class III regions.";
RL   Mamm. Genome 12:916-924(2001).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 3).
RC   TISSUE=Salivary gland;
RA   Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA   Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 5-1210 FROM N.A. (ISOFORM 1).
RC   TISSUE=Muscle, and Uterus;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE OF 195-1210 FROM N.A., AND CHARACTERIZATION.
RC   TISSUE=Histiocytic lymphoma;
RX   MEDLINE=93207535; PubMed=8457211;
RA   Milner C.M., Campbell R.D.;
RT   "The G9a gene in the human major histocompatibility complex encodes a
RT   novel protein containing ankyrin-like repeats.";
RL   Biochem. J. 290:811-818(1993).
RN   [7]
RP   ENZYMATIC ACTIVITY, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21326082; PubMed=11316813;
RA   Tachibana M., Sugimoto K., Fukushima T., Shinkai Y.;
RT   "Set domain-containing protein, G9a, is a novel lysine-preferring
RT   mammalian histone methyltransferase with hyperactivity and specific
RT   selectivity to lysines 9 and 27 of histone H3.";
RL   J. Biol. Chem. 276:25309-25317(2001).
RN   [8]
RP   IDENTIFICATION IN COMPLEX WITH E2F6; TFDP1; MAX; MGA; EUHMTASE1; CBX3;
RP   RING1; RNF2; MBLR; L3MBTL2 AND YAF2.
RX   MEDLINE=21999559; PubMed=12004135;
RA   Ogawa H., Ishiguro K.-I., Gaubatz S., Livingston D.M., Nakatani Y.;
RT   "A complex with chromatin modifiers that occupies E2F- and
RT   Myc-responsive genes in G0 cells.";
RL   Science 296:1132-1136(2002).
CC   -!- FUNCTION: Histone methyltransferase. Preferentially methylates
CC       Lys-9 of histone H3 and Lys-27 of histone H3 (in vitro). H3 Lys-9
CC       methylation represents a specific tag for epigenetic
CC       transcriptional repression by recruiting HP1 proteins to
CC       methylated histones. Probably targeted to histone H3 by different
CC       DNA-binding proteins like E2F6, MGA, MAX and/or DP1. Also
CC       methylates histone H1 (By similarity).
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + histone L-lysine =
CC       S-adenosyl-L-homocysteine + histone N(6)-methyl-L-lysine.
CC   -!- SUBUNIT: Part of the E2F6.com-1 complex in G0 phase composed of
CC       E2F6, MGA, MAX, TFDP1, CBX3, BAT8, EUHMTASE1, RING1, RNF2, MBLR,
CC       L3MBTL2 and YAF2.
CC   -!- SUBCELLULAR LOCATION: Nuclear; associates with euchromatic
CC       regions. Does not associate with heterochromatin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=Q96KQ7-1; Sequence=Displayed;
CC       Name=2; Synonyms=NG36G9a-SPI;
CC         IsoId=Q96KQ7-2; Sequence=VSP_002211;
CC       Name=3; Synonyms=NG36;
CC         IsoId=Q96KQ7-3; Sequence=VSP_002212, VSP_002213;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues examined, with high
CC       levels in fetal liver, thymus, lymph node, spleen and peripheral
CC       blood leukocytes and lower level in bone marrow.
CC   -!- SIMILARITY: Belongs to the histone-lysine methyltransferase
CC       family. Suvar3-9 subfamily.
CC   -!- SIMILARITY: Contains 7 ANK repeats.
CC   -!- SIMILARITY: Contains 1 pre-SET domain.
CC   -!- SIMILARITY: Contains 1 SET domain.
CC   -!- SIMILARITY: Contains 1 post-SET domain.
CC   -!- CAUTION: NG36 and G9a were originally thought to derive from two
CC       separate genes. Ref.3 shows that all G9A transcripts also contain
CC       the in frame coding sequence of NG36.
CC   -!- CAUTION: Ref.1 and Ref.2 sequences differ from that shown due to
CC       erroneous gene model prediction.
CC   -!- CAUTION: It is not known whether Met-1 or Met-21 is the initiator
CC       methionine.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF134726; AAD21811.1; ALT_SEQ.
DR   EMBL; AF134726; AAD21812.1; ALT_SEQ.
DR   EMBL; AP000502; BAB63294.1; ALT_SEQ.
DR   EMBL; AP000502; BAB63295.1; ALT_SEQ.
DR   EMBL; AJ315532; CAC86666.1; -.
DR   EMBL; AK056936; BAB71314.1; -.
DR   EMBL; BC002686; AAH02686.1; ALT_INIT.
DR   EMBL; BC018718; AAH18718.1; ALT_INIT.
DR   EMBL; BC020970; AAH20970.1; ALT_INIT.
DR   EMBL; X69838; CAA49491.1; ALT_INIT.
DR   Genew; HGNC:14129; BAT8.
DR   MIM; 604599; -.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR003616; PostSET.
DR   InterPro; IPR007728; Pre-SET.
DR   InterPro; IPR001214; SET.
DR   InterPro; IPR003606; Zn2-binding.
DR   Pfam; PF00023; ank; 6.
DR   Pfam; PF05033; Pre-SET; 1.
DR   Pfam; PF00856; SET; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 6.
DR   SMART; SM00468; PreSET; 1.
DR   SMART; SM00317; SET; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 5.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
```

```
DR   PROSITE; PS50868; POST_SET; FALSE_NEG.
DR   PROSITE; PS50867; PRE_SET; 1.
DR   PROSITE; PS50280; SET; 1.
KW   Transferase; Methyltransferase; Chromatin regulator; Nuclear protein;
KW   ANK repeat; Repeat; Alternative splicing.
FT   DOMAIN        2     13       POLY-ALA.
FT   DOMAIN      160    163       POLY-ALA.
FT   DOMAIN      300    326       POLY-GLU.
FT   REPEAT      649    678       ANK 1.
FT   REPEAT      684    713       ANK 2.
FT   REPEAT      717    746       ANK 3.
FT   REPEAT      750    780       ANK 4.
FT   REPEAT      784    813       ANK 5.
FT   REPEAT      817    846       ANK 6.
FT   REPEAT      850    879       ANK 7.
FT   DOMAIN      972   1035       PRE-SET.
FT   DOMAIN     1037   1159       SET.
FT   DOMAIN     1164   1180       POST-SET.
FT   VARSPLIC    373    406       Missing (in isoform 2).
FT                                /FTId=VSP_002211.
FT   VARSPLIC    195    202       PPVPEKRP -> VSGMGEMG (in isoform 3).
FT                                /FTId=VSP_002212.
FT   VARSPLIC    203   1210       Missing (in isoform 3).
FT                                /FTId=VSP_002213.
FT   CONFLICT      5      5       A -> Q (IN REF. 5; AAH18718).
FT   CONFLICT     13     13       A -> G (IN REF. 5; AAH20970).
FT   CONFLICT     55     55       N -> T (IN REF. 1 AND 4).
FT   CONFLICT    168    169       PP -> TR (IN REF. 2; AAH02686).
FT   CONFLICT    178    178       P -> S (IN REF. 3).
FT   CONFLICT    985    985       C -> R (IN REF. 3).
FT   CONFLICT    994    994       C -> R (IN REF. 6).
SQ   SEQUENCE   1210 AA;  132383 MW;  E8508923CA66911A CRC64;

  Query Match             4.9%;  Score 282.5;  DB 1;  Length 1210;
  Best Local Similarity   23.3%;  Pred. No. 3.8e-09;
  Matches  127;  Conservative   85;  Mismatches  201;  Indels  133;  Gaps   18;

Qy          7 VVEYLLQN-GASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEA 65
              |:  || |   : |:       ||| |   |  |: ::||: ||: || |    ||| ||
Db        666 VILMLLDNLDPNFQSDQQSKRTPLHAAAQKGSVEICHVLLQAGANINAVDKQQRTPLMEA 725

Qy         66 AIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEE 125
               :   ::|   ::| |      :  || | |                      :|: || |
Db        726 VVNNHLEVARYMVQRGGCVYSKEEDGSTCLH-------------------HAAKIGNLE 765

Qy        126 KMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNA 185
              :  ||:   |: :| |    ||:  || :  :::::||  ||||   |  : : || |
Db        766 MVSLLLSTGQVDVNAQDSGGWTPIIWAAEHKHIEVIRMLLTRGADVTLTDNEENICLHWA 825

Qy        186 CSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNK 245
                |   : |:|:     ::|::    |||| || ::  :  | || ||:| | |
Db        826 SFTGSAAIAEVLLNARCDLHAVNYHGDTPLHIAARESYHDCVLLFLSRGANPELRNKEGD 885

Qy        246 SAIDLAP-------TPQLKERLAYEFKGHSLLQAARE--ADVTRIKKHLSLEMVN----- 291
              :| || |          ||  :|     |:  ::  :    || |  ::: :  ||
Db        886 TAWDLTPERSDVWFALQLNRKLRLGV-GNRAIRTEKIICRDVARGYENVPIPCVNGVDGE 944

Qy        292 -----FKH-PQTHETAL-----------HCAAASPYPKRKQIC-ELLLR----KGANINE 329
                  :|:  :  ||:             ||            :| :| :|     |   : :
Db        945 PCPEDYKYISENCETSTMNIDRNITHLQHCTCVDDCSSSNCLCGQLSIRCWYDKDGRLLQ 1004

Qy        330 KTKEFLTPLHVASEKA-------HNDVV------------------------------ 350
              : :   ||    :|         | ||
Db       1005 EFNKIEPPLIFECNQACSCWRNCKNRVVQSGIKVRLQLYRTAKMGWGVRALQTIPQGTFI 1064

Qy        351 -----EVVVKHEAKVNA-------LDNL-GQTSLHRAAYCGHLQTCRLLLSYGCDPNIIS 397
                  |::   || |          |||  |:      | | |::       ::: ||||||
Db       1065 CEYVGELISDAEADVREDDSYLFDLDNKDGEVYCIDARYYGNISR---FINHLCDPNIIP 1121

Qy        398 LQGFTALQMGNENVQ---------QLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTV 448
              :: |    | :::::         : :: |  ||    ||         |:::    |
Db       1122 VRVF----MLHQDLRFPRIAFFSSRDIRTGEELGFDYGDR-------FWDIKSKYFTCQC 1170
```

```
Qy         449 QSVNCR 454
                 |  |:
Db        1171 GSEKCK 1176
```

Search completed: March 25, 2004, 16:42:18
Job time : 41.1624 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:37:35 ; Search time 44.188 Seconds
(without alignments)
7854.389 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

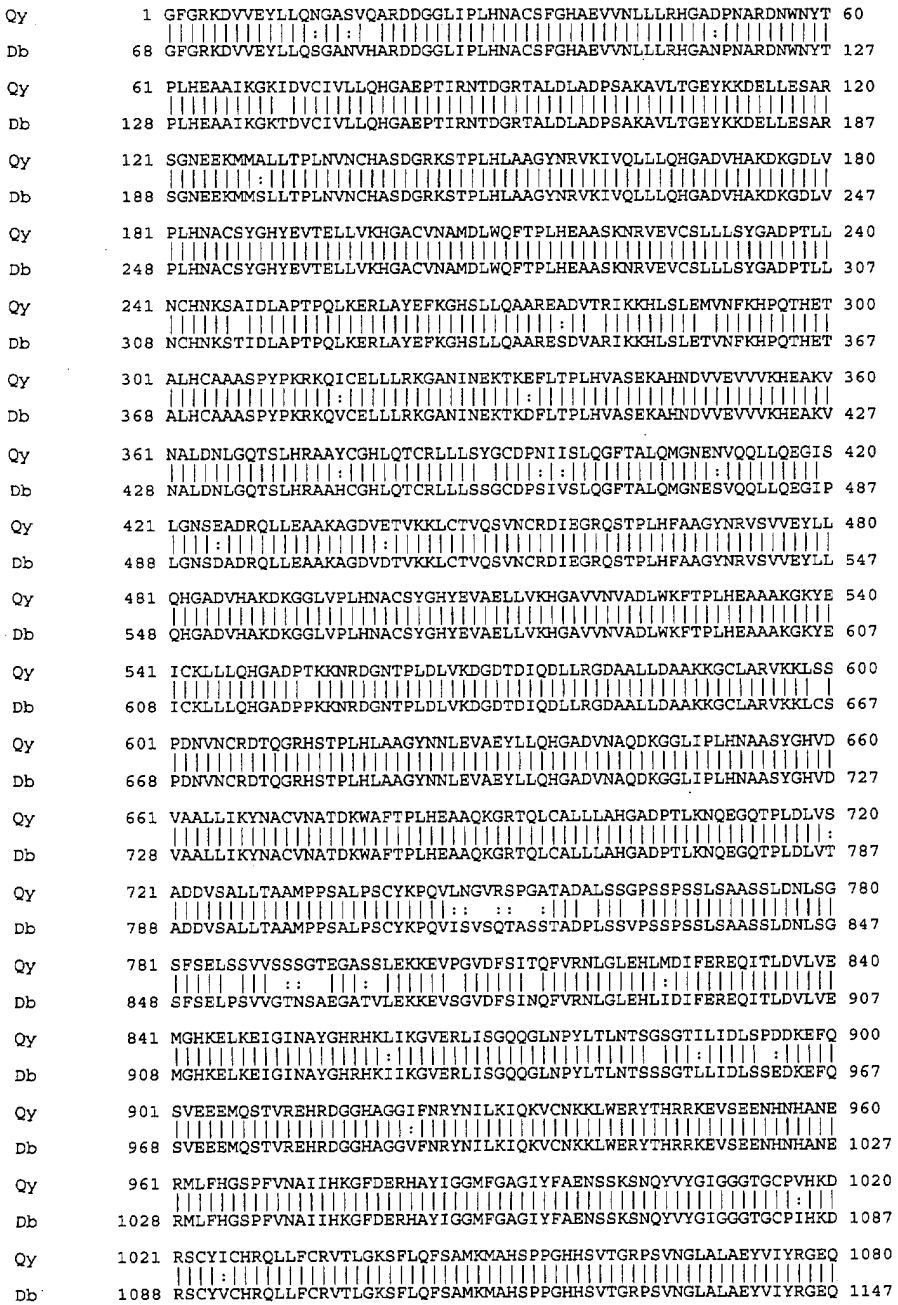
SUMMARIES

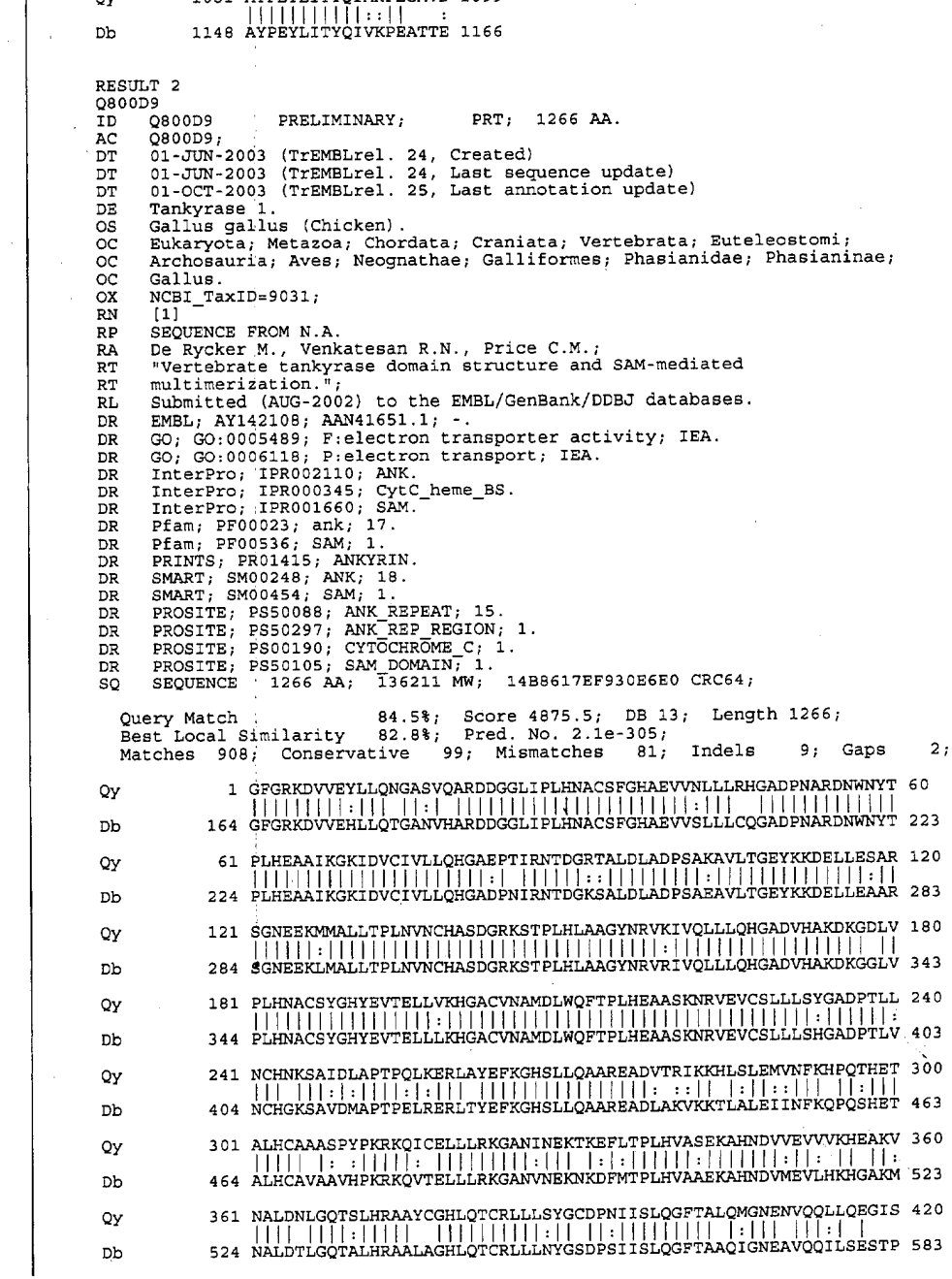
Result No.	Score	% Query Match	Length	DB	ID	Description
1	5501	95.4	1167	13	Q800E0	Q800e0 gallus gall
2	4875.5	84.5	1266	13	Q800D9	Q800d9 gallus gall
3	4285.5	74.3	976	11	Q8BX62	Q8bx62 mus musculu
4	3835.5	66.5	1181	5	Q9VBP3	Q9vbp3 drosophila
5	3835.5	66.5	1181	5	Q9XZ37	Q9xz37 drosophila
6	2727	47.3	600	11	Q8BXH7	Q8bxh7 mus musculu
7	876.5	15.2	1887	4	Q7Z3G4	Q7z3g4 homo sapien
8	875.5	15.2	1726	11	Q8VC68	Q8vc68 mus musculu
9	873.5	15.1	1943	11	Q61307	Q61307 mus musculu
10	867.5	15.0	1136	6	Q9N180	Q9n180 bos taurus
11	864.5	15.0	2622	11	O70511	O70511 rattus norv
12	864	15.0	1863	4	Q7Z3L5	Q7z3l5 homo sapien
13	863.5	15.0	1762	11	O88521	O88521 rattus norv
14	854.5	14.8	1219	11	Q8C8R3	Q8c8r3 mus musculu
15	842	14.6	1719	4	Q13768	Q13768 homo sapien
16	841.5	14.6	843	11	P97582	P97582 rattus norv
17	840.5	14.6	1856	4	Q99407	Q99407 homo sapien
18	839.5	14.6	1848	11	Q61302	Q61302 mus musculu
19	836	14.5	1549	5	Q9V4B1	Q9v4b1 drosophila
20	834	14.5	1549	5	Q24241	Q24241 drosophila
21	820	14.2	1159	5	Q9NCP8	Q9ncp8 drosophila
22	814.5	14.1	1809	5	Q17487	Q17487 caenorhabdi
23	814.5	14.1	1841	5	Q8MQG0	Q8mqg0 caenorhabdi
24	814.5	14.1	1867	5	Q17486	Q17486 caenorhabdi
25	814.5	14.1	2039	5	Q17489	Q17489 caenorhabdi
26	810.5	14.0	1815	5	Q17488	Q17488 caenorhabdi
27	810.5	14.0	6994	5	Q17343	Q17343 caenorhabdi
28	810.5	14.0	6994	5	Q17490	Q17490 caenorhabdi
29	807.5	14.0	1786	5	Q17344	Q17344 caenorhabdi
30	793	13.7	1009	5	Q8SWY2	Q8swy2 drosophila
31	678	11.8	919	4	Q8N8A2	Q8n8a2 homo sapien
32	669	11.6	1076	11	Q8BTI7	Q8bti7 mus musculu
33	640	11.1	1179	13	Q8AX92	Q8ax92 fugu rubrip
34	631	10.9	1614	13	Q7T1G6	Q7t1g6 brachydanio
35	629.5	10.9	1088	4	Q13484	Q13484 homo sapien
36	628	10.9	791	4	Q9UFA4	Q9ufa4 homo sapien
37	606	10.5	638	4	Q8IZ72	Q8iz72 homo sapien
38	601	10.4	1913	5	Q9GRV5	Q9grv5 caenorhabdi
39	584.5	10.1	1619	5	Q9VMR4	Q9vmr4 drosophila
40	554.5	9.6	1098	11	Q61304	Q61304 mus musculu
41	551.5	9.6	531	11	Q8C036	Q8c036 mus musculu
42	545.5	9.5	1411	5	Q02989	Q02989 latrodectus
43	538	9.3	2486	4	O75179	O75179 homo sapien
44	534.5	9.3	1599	11	Q99NH0	Q99nh0 mus musculu
45	527	9.1	743	4	Q8TB46	Q8tb46 homo sapien

ALIGNMENTS

RESULT 1
Q800E0
ID Q800E0 PRELIMINARY; PRT; 1167 AA.
AC Q800E0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tankyrase 2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA De Rycker M., Venkatesan R.N., Price C.M.;
RT "Vertebrate tankyrase domain structure and SAM-mediated
RT multimerization.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY142107; AAN41650.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 16.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 16.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 15.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 1167 AA; 126931 MW; 7842BF482B9229E9 CRC64;

Query Match 95.4%; Score 5501; DB 13; Length 1167;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1040; Conservative 32; Mismatches 27; Indels 0; Gaps 0;





RESULT 2
Q800D9
ID Q800D9 PRELIMINARY; PRT; 1266 AA.
AC Q800D9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tankyrase 1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA De Rycker M., Venkatesan R.N., Price C.M.;
RT "Vertebrate tankyrase domain structure and SAM-mediated
RT multimerization.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY142108; AAN41651.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 17.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 18.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 15.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
SQ SEQUENCE 1266 AA; 136211 MW; 14B8617EF930E6E0 CRC64;

Query Match 84.5%; Score 4875.5; DB 13; Length 1266;
Best Local Similarity 82.8%; Pred. No. 2.1e-305;
Matches 908; Conservative 99; Mismatches 81; Indels 9; Gaps 2;

```
Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
Db        584 VRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 643

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
Db        644 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 703

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
Db        704 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLRGDAALLDAAKKGCLARVQKLCT 763

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
Db        764 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 823

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
Db        824 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 883

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
Db        884 ADDIRALLIDAMPPEALPTCFKLQAT-------VVSASLISPASTPSCLSAASSIDNLTG 936

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
Db        937 PLAELAVGGASNTGDGAAGTERKEGEVSGLDMNITQFLKSLGLEHLRDIFETEQITLDVL 996

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
Db        997 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVSQGTILLDLAPDDKE 1056

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
Db       1057 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1116

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
Db       1117 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1176

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
Db       1177 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1236

Qy       1079 EQAYPEYLITYQIMRPE 1095
Db       1237 EQAYPEYLITYQIVKPE 1253
```

```
RESULT 3
Q8BX62
ID   Q8BX62      PRELIMINARY;      PRT;   976 AA.
AC   Q8BX62;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TRF1-interacting ankyrin-related ADP-ribose polymerase homolog.
GN   TNKS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK048860; BAC33475.1; -.
DR   PIR; PT0551; PT0551.
DR   MGD; MGI:1341087; Tnks.
DR   GO; GO:0005489; F:electron transporter activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 15.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 13.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 12.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
SQ   SEQUENCE   976 AA;  106663 MW;  356B10CE9C33ECF1 CRC64;

  Query Match             74.3%;  Score 4285.5;  DB 11;  Length 976;
  Best Local Similarity   82.0%;  Pred. No. 1.6e-267;
  Matches  798;  Conservative   88;  Mismatches   72;  Indels   15;  Gaps    3;

Qy        128 MALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACS 187
Db          1 MALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVPLHNACS 60

Qy        188 YGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSA 247
Db         61 YGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVNCHGKSA 120

Qy        248 IDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAA 307
Db        121 VDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETALHCAVA 180

Qy        308 SPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLG 367
Db        181 SLHPKRKQVAELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKMNALDSLG 240

Qy        368 QTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEAD 427
Db        241 QTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTPMRTSDVD 300

Qy        428 RQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVH 487
Db        301 YRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLLHHGADVH 360

Qy        488 AKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQ 547
Db        361 AKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYEICKLLLK 420

Qy        548 HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCR 607
Db        421 HGADPTKKNRDGNTPLDLVKEGDTDIQDLLRGDAALLDAAKKGCLARVQKLCTPENINCR 480

Qy        608 DTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 667
Db        481 DTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVDIAALLIK 540

Qy        668 YNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSAL 727
Db        541 YNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLATADDIRAL 600

Qy        728 LTAAMPPSALPSCYKPQ---VLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSE 784
Db        601 LIDAMPPEALPTCFKPQATVVSASVISPAST----------PSCLSAASSIDNLTGPLTD 650

Qy        785 LSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMG 842
Db        651 LAVGGASNAGDGAAGAERKEGEVAGLDMNIIQFLKSLGLEHLRDIFETEQITLDVLADMG 710

Qy        843 HKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSV 902
Db        711 HEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKEYQSV 770
```

```
Qy        903 EEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERM 962
              |||||||:|||||||:|||||||||:::|||| |||| ||: ||:|||||||||| ||||
Db        771 EEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHHNERM 830

Qy        963 LFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS 1022
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        831 LFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRS 890

Qy       1023 CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAY 1082
              |||||||:|||||||||||||||| |||||:||||||| |||||||||| ||||||||||
Db        891 CYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAY 950

Qy       1083 PEYLITYQIMRPE 1095
              |||||||||||:||
Db        951 PEYLITYQIMKPE 963
```

```
RESULT 4
Q9VBP3
ID   Q9VBP3      PRELIMINARY;      PRT;  1181 AA.
AC   Q9VBP3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   BCDNA:LD22548 protein.
GN   BCDNA:LD22548 OR CG4719 OR CG17487.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
```

```
DR   EMBL; AE003752; AAF56487.1; -.
DR   HSSP; Q00420; 1AWC.
DR   FlyBase; FBgn0027508; BcDNA:LD22548.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR001660; SAM.
DR   Pfam; PF00023; ank; 17.
DR   Pfam; PF00536; SAM; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   SMART; SM00454; SAM; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 14.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50105; SAM_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1181 AA;  127944 MW;  EB6F0F94E6A3D504 CRC64;
```

Query Match 66.5%; Score 3835.5; DB 5; Length 1181;
Best Local Similarity 65.3%; Pred. No. 2.3e-238;
Matches 740; Conservative 156; Mismatches 182; Indels 55; Gaps 12;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |:||::|||:|| :|||:|| |:||| |||| ||||||||| |||: || || ||||||
Db         66 GYGRREVVEFLLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYT 125

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||| |||:|||: |||||| ||||:: :| |:||| : : ||||||:|||||:||
Db        126 PLHEAASKGKVDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAAR 185

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              || |::::|||||||||||||||:|||||||||||: ||::|| :|||||||||| ||
Db        186 SGAEDRLLALLTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLV 245

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||::||:||:: || ||| ||| ||||||||||:|||||||||| |||||||
Db        246 PLHNACSYGHFDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLL 305

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              |||:||||| ||| :|:||:|:|:||| || | |: ||:| || :  |:||| || | :|
Db        306 NCHSKSAIDAAPTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDT 365

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              || |  ||  ||||: ||| |||: :||| | ||||||:|:|  | | :|||:|  |||
Db        366 PLHLAVVSPDGKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKV 425

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||:|||| ||| |      |  |||||  | ||:||:|  || |: :::| :|||:
Db        426 NALDSLGQTPLHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK---- 479

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                  :::  |||||||||::||:::      |||||||:|| |||||||||:||| ||::
Db        480 -NPPDSETHLLEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQF 538

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
              ||:|||:|:| |||||||||||||||||| |||||||| |||:|||||||||||||||||
Db        539 LLEHGAEVYAADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGK 598

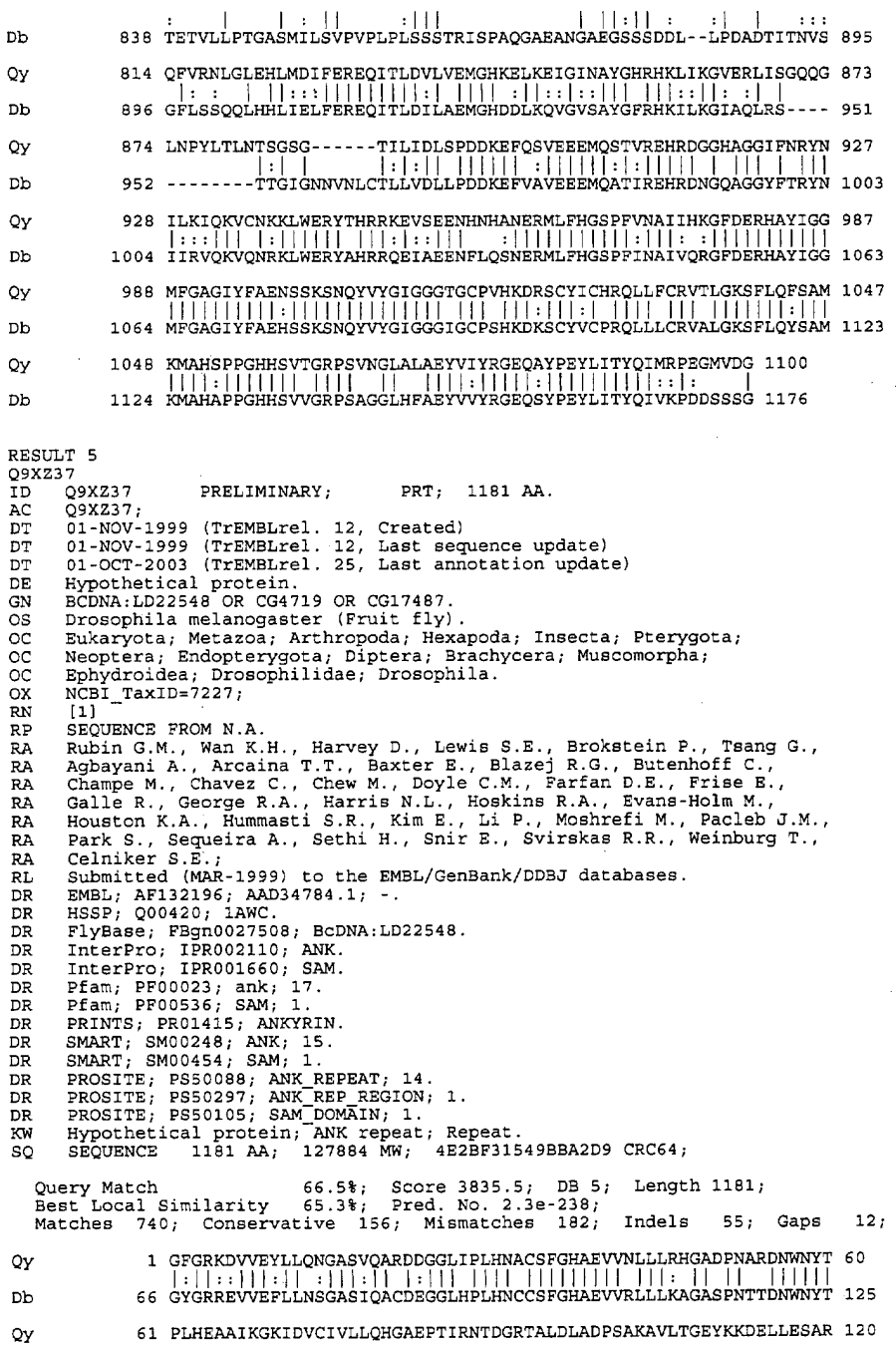
Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
              |:||||||:||||| |||||| || |||| | |: :|||| :|||||||| ||||::|
Db        599 YDICKLLLKHGADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRL 658

Qy        599 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658
              :|:::|||| |||:|||||||||||| | |||||::|||||||||||||||:||||
Db        659 VTPESINCRDAQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGH 718

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
              :|:|||||:   |||||| |||||||||||||||:|||||||  :|||||||::|
Db        719 LDIAALLIKHKTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIEL 778

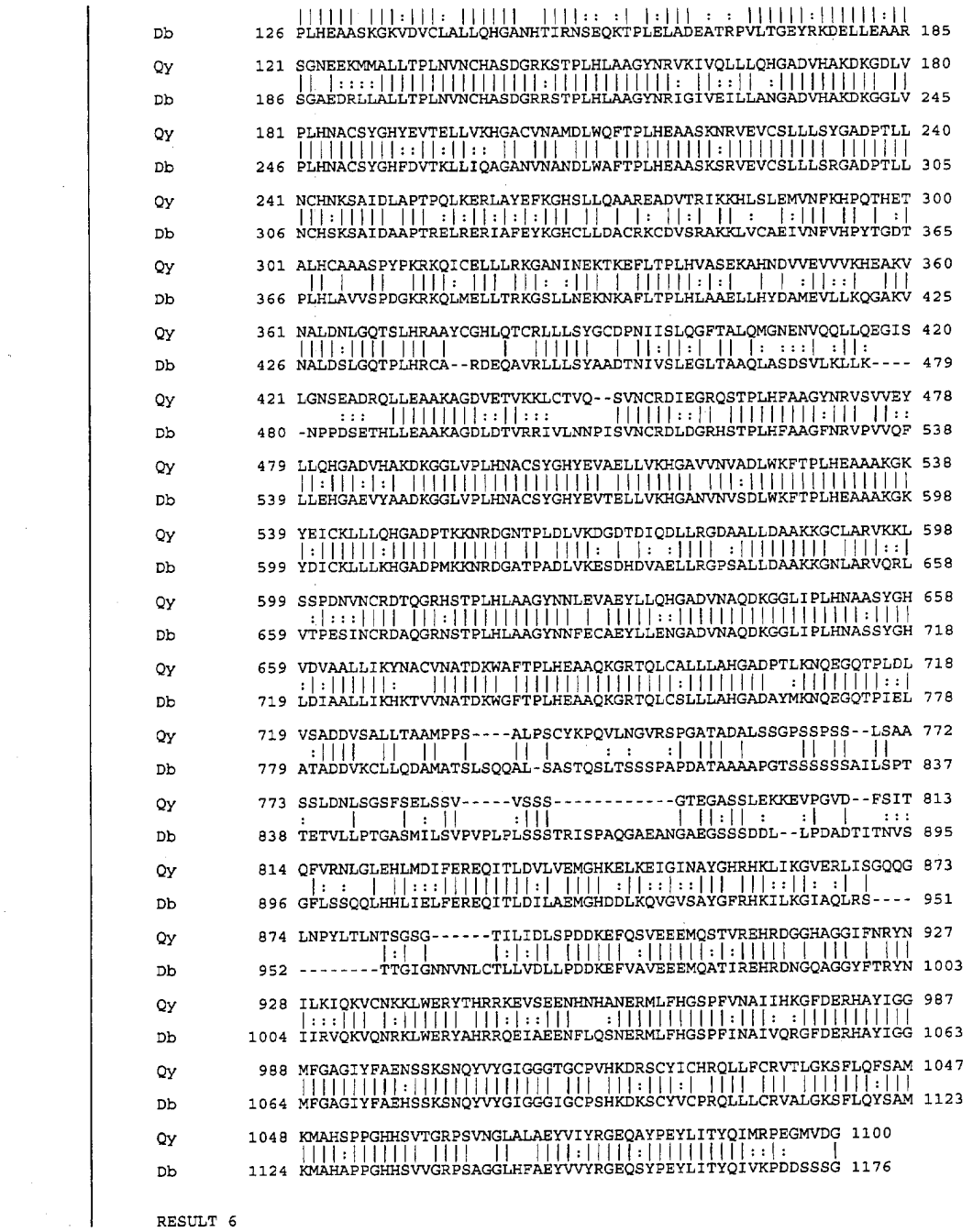
Qy        719 VSADDVSALLTAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAA 772
              :|||| || || |  || |      : :  :| ||| |     || || ||
Db        779 ATADDVKCLLQDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPT 837

Qy        773 SSLDNLSGSFSELSSV-----VSSS-----------GTEGASSLEKKEVPGVD--FSIT 813
```



RESULT 5
Q9XZ37
ID Q9XZ37 PRELIMINARY; PRT; 1181 AA.
AC Q9XZ37;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN BCDNA:LD22548 OR CG4719 OR CG17487.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF132196; AAD34784.1; -.
DR HSSP; Q00420; 1AWC.
DR FlyBase; FBgn0027508; BcDNA:LD22548.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001660; SAM.
DR Pfam; PF00023; ank; 17.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 15.
DR SMART; SM00454; SAM; 1.
DR PROSITE; PS50088; ANK_REPEAT; 14.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50105; SAM_DOMAIN; 1.
KW Hypothetical protein; ANK repeat; Repeat.
SQ SEQUENCE 1181 AA; 127884 MW; 4E2BF31549BBA2D9 CRC64;

Query Match 66.5%; Score 3835.5; DB 5; Length 1181;
Best Local Similarity 65.3%; Pred. No. 2.3e-238;
Matches 740; Conservative 156; Mismatches 182; Indels 55; Gaps 12;



RESULT 6

```
Q8BXH7
ID   Q8BXH7      PRELIMINARY;      PRT;   600 AA.
AC   Q8BXH7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   TANKYRASE-related protein (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK047094; BAC32960.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 14.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 11.
DR   PROSITE; PS50088; ANK_REPEAT; 11.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
FT   NON_TER     600    600
SQ   SEQUENCE   600 AA;  65307 MW;  2F667AD865F807F6 CRC64;

  Query Match           47.3%;  Score 2727;  DB 11;  Length 600;
  Best Local Similarity  97.0%;  Pred. No. 2.5e-167;
  Matches  518;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||||:|||||||||||||||||||||||||||:||| ||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLQHGAHPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||:||:|||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGDYKEDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              |||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAEYNRVKIVQLLLHHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRIEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              |||||||||||| |||||:|||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTAQLKERLSYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              |||||||||||||||||||||||| |||||||||||||| ||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANTNEKTKEFLTPLHVASENAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||:|||||||||:|||||||||||||||||||||||||||||||||||||||||||| |
Db        427 NALDSLGQTSLHRAAHCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGAS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGHSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 534
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 600


RESULT 7
Q7Z3G4
ID   Q7Z3G4      PRELIMINARY;      PRT;  1887 AA.
AC   Q7Z3G4;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein DKFZp686P17114 (Fragment).
GN   DKFZP686P17114.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Human fetal kidney;
RA   Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX537917; CAD97900.1; -.
KW   Hypothetical protein.
FT   NON_TER     1      1
SQ   SEQUENCE   1887 AA;  205502 MW;  CAE428F76E8C5B5B CRC64;

  Query Match           15.2%;  Score 876.5;  DB 4;  Length 1887;
  Best Local Similarity  30.5%;  Pred. No. 3.4e-47;
  Matches  273;  Conservative  126;  Mismatches  362;  Indels  135;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              | :|| ||| |:| | |  | || | | |||| :|: :||: ||:   :|||
Db        105 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 164

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :||  ::   || | | :|     |     :|  : |   :
Db        165 YMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 224

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              | |  :: |  |||  : |         ||||:|| |  : :  |||   | |
Db        225 ALHIAARKDDTKAAALLPQNDNNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 284

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |  |: : :||:  || ::|     |||| |     :|  :||
Db        285 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRA 344

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |  |                | | | : |     ||:  :  :|
Db        345 A-PVLSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 376

Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | || | |     ||||:| :|
Db        377 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTPLHIACKKNRI 433

Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
              |:|:::|| | : |:    | | :| ||: ||:      |: :|  ||  :::| ||| |
Db        434 KVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRGETALHMA 493

Qy        408 -----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVNCRDIEGR 459
                  | |: |:|:|  :      | |  |  :|: |  : |::|     | |     |
Db        494 ARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGTSPNAATTSG- 552

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
              ||||  :|      | :|| ||| :    | |  ||| |  ||  ||| ||::  |
Db        553 -YTPLHLSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVANLLLQKSASP 611

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL- 577
              : |     |||| ||      ::  |||  || |     ::| ||| :  |     ||
Db        612 DAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATTLL 671

Qy        578 ----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNN 625
                       :| |::  ||::| :  |  |   : |||  :  |   |||||||  :
Db        672 EYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNANVNLSNKSG--LTPLHLAAQEDR 729

Qy        626 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
              : ||| |:  || |:|| | |  |||     ||:: :   |:::| ||| |  :||||:
Db        730 VNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTPLHQ 789
```

```
Qy        686 AAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLT-AAMPPSALP 738
              |||:| | :  :|| : | |       | | | :      :|  |   ::|   |  : :
Db        790 AAQQGHTHIINVLLQNNASPNELTVNGNTALGIARRLGYISVVDTLKIVTEETMTTTTVT 849

Qy        739 SCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 795
                :|   |: :| |               |   :  |:: : ||
Db        850 EKHKMNVPETMNEVLD-----------MSDDEVRKANAPEMLSDG-------------- 883

Qy        796 GASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ-ITLDVLVEMGHKELKEIG 850
                                      |:: |: | |  :| |    :| ::|||:|
Db        884 -------------------------EYISDVEEGEDAMTGDTDKYLGPQDLKELG 913


RESULT 8
Q8VC68
ID   Q8VC68      PRELIMINARY;      PRT;  1726 AA.
AC   Q8VC68;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   ANK3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC021657; AAH21657.1; -.
DR   MGD; MGI:88026; Ank3.
DR   GO; GO:0045202; C:synaptic junction; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   1726 AA;  188239 MW;  430148B67E00B1CA CRC64;

  Query Match           15.2%;  Score 875.5;  DB 11;  Length 1726;
  Best Local Similarity   28.7%;  Pred. No. 3.5e-47;
  Matches  320;  Conservative  158;  Mismatches  465;  Indels  171;  Gaps   35;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:| |     |   || |   | |||| :|: :||: ||:    :|||
Db         68 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 127

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :|| ::   || | | :|     |     :|  : |   :
Db        128 YMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 187

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
                |   |   :: |  |||   : |          ||||:|| |  : :  |||   | |
Db        188 ALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 247

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |   |:  : :||:  || ::|       ||||  |     :|  :||
Db        248 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRS 307

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |   |                    | | |  | : |      ||:   :  :|

Db        308 A-PILSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 339

Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | |: | |     ||||:| :|
Db        340 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKASPNAKALNGFTPLHIACKKNRI 396

Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
              |:|:::|| | : |:   | | :| ||: ||:      |: :|  ||  :::| ||| |
Db        397 RVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRGETALHMA 456

Qy        408 -----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVNCRDIEGR 459
                  | |: |:|:|  :     | |  |  :|: |  : |::|     | |     |
Db        457 ARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGASPNAATTSG- 515

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
              :|||| ||       |  :|| ||| :    | |  ||| |  ||  ||| ||::  |
Db        516 -YTPLHLAAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVASLLLQKSASP 574

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL- 577
              : |    |||| ||     ::  ||| || |    ::| ||| :  |    ||   |
Db        575 DAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATSLL 634

Qy        578 -----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNN 625
                        :| |::  ||::| :  |  | | : |||  :  |   |||||||  :
Db        635 EYGADANAVTRQGIASVHLAAQEGHVDMVSLLLSRNANVNLSNKSG--LTPLHLAAQEDR 692

Qy        626 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
              : ||| |:  || |:|| | |  |     ||:: :   |::::| ||| |  :| ||:
Db        693 VNVAEVLVNQGAHVDAQTKMGYTPPHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTALHQ 752

Qy        686 AAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS-------ADDVSALLTAAMPPSALP 738
              |||:| | :  :|| : | |       | | | :           | :  :    | : :
Db        753 AAQQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVDTLKVVTEEIMTTTTIT 812

Qy        739 SCYK---PQVLNGV--RSPGATADALSS------GP---------SSPSSLSAASSLDNL 778
                :|   |: :| |   |          ||::      ||          | |:     ||
Db        813 EKHKMNVPETMNEVLDMSDDEGEDAITGDTDKYLGPQDLKELGDDSLPAEGYVGFSLGAR 872

Qy        779 SGSFSELSSVVS-----SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQI 833
              | |    ||  |     ||    :  :|:  ||  :  :| | |    : |  :
Db        873 SASLRSFSSDRSYTLNRSSYARDSMMIEELLVPSKEQHLT-FTREFDSDSLRH-YSWAAD 930

Qy        834 TLDVL------VEMGHKELKEIGINAYG--------HRHKLI----------KGVERLIS 869
              |||  :      |  |   |    ::| |         | ::|           :   ||:
Db        931 TLDNVNLVSSPVHSGF--LVSFMVDARGGSMRGSRHHGMRIIIPPRKCTAPTRITCRLVK 988

Qy        870 GQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS-VEEEMQSTVREHRDGGHAGGIFNRYNI 928
                :  ||   :   | :|::: |   :|   |  |:           | | :  :
Db        989 RHKLANPPPMVEGEGLASRLVEMGPAGAQFLGPVIVEIP----------HFGSMRGKERE 1038

Qy        929 LKIQKVCNKKLWE--RYTHRRKEVSE-----ENHNHANERMLFHGSPFVNAIIHKGFDER 981
              | :  | : |:  ::  : ::::|     :   :  |  :   |:  :  || | | :
Db       1039 LIVLRSENGETWKEHQFDSKNEDLAELLNGMDEELDSPEEL---GTKRICRIITKDFPQ- 1094

Qy        982 HAYIGGMFGAGIYFAENS---SKSNQYVYGIGGG 1012
                        |||  |   :|||   |  ||
Db       1095 ------------YFAVVSRIKQESNQ--IGPEGG 1114


RESULT 9
Q61307
ID   Q61307      PRELIMINARY;      PRT;  1943 AA.
AC   Q61307; Q61305; Q61306; Q61308; Q61309; Q61310; O08866; O08867;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ankyrin 3 (Ankyrin G) (Epithelial ankyrin) (Ankyrin-3).
GN   ANK3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
```

```
RP   SEQUENCE FROM N.A. (ISOFORMS 1-6).
RC   STRAIN=C57BL/6J; TISSUE=Kidney;
RX   MEDLINE=95340633; PubMed=7615634;
RA   Peters L.L., John K.M., Lu F.M., Eicher E.M., Higgins A., Yialamas M.,
RA   Turtzo L.C., Otsuka A.J., Lux S.E.;
RT   "Ank3 (epithelial ankyrin), a widely distributed new member of the
RT   ankyrin gene family and the major ankyrin in kidney, is expressed in
RT   alternatively spliced forms, including forms that lack the repeat
RT   domain.";
RL   J. Cell Biol. 130:313-330(1995).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 7 AND 8).
RC   STRAIN=C3H/HEJ; TISSUE=Bone Marrow;
RX   MEDLINE=97213781; PubMed=9060470;
RA   Hoock T.C., Peters L.L., Lux S.E.;
RT   "Isoforms of ankyrin-3 that lack the NH2-terminal repeats associate
RT   with mouse macrophage lysosomes.";
RL   J. Cell Biol. 136:1059-1070(1997).
CC   -!- FUNCTION: ANKYRINS ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTO-
CC       SKELETAL ELEMENTS; THEY BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN
CC       BAND 4.2, TO NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85,
CC       AND TO THE CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND
CC       DESMIN. ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO
CC       THE CYTOPLASMIC DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;
CC       THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.
CC   -!- SUBCELLULAR LOCATION: PLASMA-MEMBRANE ASSOCIATED. ISOFORMS 7 AND 8
CC       ARE ASSOCIATED WITH THE LYSOSOMAL MEMBRANE.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=8;
CC         Comment=Additional isoforms seem to exist;
CC       Name=1;
CC         IsoId=Q61307-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q61307-2; Sequence=VSP_050350, VSP_050353;
CC       Name=3;
CC         IsoId=Q61307-3; Sequence=VSP_050350, VSP_050353, VSP_050355;
CC       Name=4;
CC         IsoId=Q61307-4; Sequence=VSP_050354, VSP_050355;
CC       Name=5;
CC         IsoId=Q61307-5; Sequence=VSP_050352;
CC       Name=6;
CC         IsoId=Q61307-6; Sequence=VSP_050354;
CC       Name=7;
CC         IsoId=Q61307-7; Sequence=VSP_050349, VSP_050351, VSP_050355;
CC       Name=8;
CC         IsoId=Q61307-8; Sequence=VSP_050349, VSP_050351;
CC   -!- TISSUE SPECIFICITY: MAJOR FORM OF ANKYRIN IN KIDNEY AND OTHER
CC       EPITHELIAL TISSUES. ALSO EXPRESSED IN MACROPHAGES, MEGAKARYOCYTES,
CC       LEYDIG CELLS, CARDIAC, SMOOTH AND SKELETAL MUSCLE, INITIAL
CC       SEGMENTS OF AXONS, AND NODES OF RANVIER.
CC   -!- SIMILARITY: CONTAINS 24 ANK REPEATS.
DR   EMBL; L40631; AAB01603.1; -.
DR   EMBL; L40632; AAB01605.1; -.
DR   EMBL; U89275; AAB58381.1; -.
DR   EMBL; L40631; AAB01602.1; -.
DR   EMBL; L40632; AAB01604.1; -.
DR   EMBL; L40632; AAB01606.1; -.
DR   EMBL; L40632; AAB01607.1; -.
DR   EMBL; U89274; AAB58380.1; -.
DR   PIR; T42713; T42713.
DR   PIR; T42714; T42714.
DR   PIR; T42715; T42715.
DR   PIR; T42716; T42716.
DR   HSSP; P55273; 1BI8.
DR   MGD; MGI:88026; Ank3.
DR   GO; GO:0045202; C:synaptic junction; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW   Multigene family.
FT   DOMAIN        1    826       ANION EXCHANGE PROTEIN BINDING DOMAIN.
FT   DOMAIN       23    810       24 X ANK MOTIF REPEATS.
FT   DOMAIN      856   1437       SPECTRIN BINDING DOMAIN.
FT   DOMAIN     1438   1943       REGULATORY DOMAIN (REGULATES BINDING OF
FT                                ANKYRIN TO SPECTRIN AND THE BAND 3
FT                                PROTEIN).
FT   REPEAT       23     55       ANK MOTIF 1.
FT   REPEAT       56     88       ANK MOTIF 2.
FT   REPEAT       89    121       ANK MOTIF 3.
FT   REPEAT      122    154       ANK MOTIF 4.
FT   REPEAT      155    183       ANK MOTIF 5.
FT   REPEAT      184    216       ANK MOTIF 6.
FT   REPEAT      217    249       ANK MOTIF 7.
FT   REPEAT      250    282       ANK MOTIF 8.
FT   REPEAT      283    315       ANK MOTIF 9.
FT   REPEAT      316    348       ANK MOTIF 10.
FT   REPEAT      349    381       ANK MOTIF 11.
FT   REPEAT      382    414       ANK MOTIF 12.
FT   REPEAT      415    447       ANK MOTIF 13.
FT   REPEAT      448    480       ANK MOTIF 14.
FT   REPEAT      481    513       ANK MOTIF 15.
FT   REPEAT      514    546       ANK MOTIF 16.
FT   REPEAT      547    579       ANK MOTIF 17.
FT   REPEAT      580    612       ANK MOTIF 18.
FT   REPEAT      613    645       ANK MOTIF 19.
FT   REPEAT      646    678       ANK MOTIF 20.
FT   REPEAT      679    711       ANK MOTIF 21.
FT   REPEAT      712    744       ANK MOTIF 22.
FT   REPEAT      745    777       ANK MOTIF 23.
FT   REPEAT      778    810       ANK MOTIF 24.
FT   VARSPLIC      1    805       Missing (in isoform 7 and isoform 8).
FT                                /FTId=VSP_050349.
FT   VARSPLIC      1    849       Missing (in isoform 2 and isoform 3).
FT                                /FTId=VSP_050350.
FT   VARSPLIC    833    853       Missing (in isoform 7 and isoform 8).
FT                                /FTId=VSP_050351.
FT   VARSPLIC    834    855       VRKASAPEKLSDGEYISDGEEG -> GDKCTWFKIPKVQEV
FT                                LVK (in isoform 5).
FT                                /FTId=VSP_050352.
FT   VARSPLIC    850    855       SDGEEG -> MALPHS (in isoform 2 and
FT                                isoform 3).
FT                                /FTId=VSP_050353.
FT   VARSPLIC    852    852       G -> GDKCTWFKIPKVQEVLVKS (in isoform 4
FT                                and isoform 6).
FT                                /FTId=VSP_050354.
FT   VARSPLIC   1569   1764       Missing (in isoform 3, isoform 4 and
FT                                isoform 7).
FT                                /FTId=VSP_050355.
SQ   SEQUENCE   1943 AA;  212013 MW;  B23C4B6F82904804 CRC64;

  Query Match           15.1%;  Score 873.5;  DB 11;  Length 1943;
  Best Local Similarity   30.6%;  Pred. No. 5.6e-47;
  Matches  274;  Conservative  122;  Mismatches  365;  Indels  135;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:| |    |   || |   | |||| :|: :||: ||:    :|||
Db         68 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 127

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :||  ::   || | | :|     |      :|  : |   :
Db        128 YMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 187

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |  |  ::  |  |||   : |       ||||:|| |  : :  |||   | |
Db        188 ALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 247
```

```
Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
Db        248 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRS 307

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
Db        308 A-PILSKTKN----------------------GLSPLHMATQGD------HLNCVQLLLQH 339

Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
Db        340 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKASPNAKALNGFTPLHIACKKNRI 396

Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
Db        397 RVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRGETALHMA 456

Qy        408 -----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVNCRDIEGR 459
Db        457 ARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGASPNAATTSG- 515

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
Db        516 -YTPLHLAAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVASLLLQKSASP 574

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL- 577
Db        575 DAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATSLL 634

Qy        578 -----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNN 625
Db        635 EYGADANAVTRQGIASVHLAAQEGHVDMVSLLLSRNANVNLSNKSG--LTPLHLAAQEDR 692

Qy        626 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
Db        693 VNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTALHQ 752

Qy        686 AAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS-------ADDVSALLTAAMPPSALP 738
Db        753 AAQQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVDTLKVVTEEIMTTTTIT 812

Qy        739 SCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLS-GSFSELSSVVSSSGT 794
Db        813 EKHKMNVPETMNEVLD-----------MSDDEVRKASAPEKLSDGEY-------ISDGE 853

Qy        795 EGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIG 850
Db        854 EG---------------------------------EDAITGDTDKYLGPQDLKELG 876
```

RESULT 10
Q9N180
ID Q9N180 PRELIMINARY; PRT; 1136 AA.
AC Q9N180;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ankyrin 1 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=breed Japanese Black; TISSUE=Bone marrow;
RA Matsumoto M., Inaba M., Koshino I., Saito D., Ono K.;
RT "Cloning of bovine erythrocyte ankyrin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF222766; AAF61702.1; -.
DR HSSP; Q00420; 1AWC.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 22.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat; Repeat.
FT NON_TER 1136 1136
SQ SEQUENCE 1136 AA; 123182 MW; 07CD8B1568ACE8C7 CRC64;

Query Match 15.0%; Score 867.5; DB 6; Length 1136;
Best Local Similarity 26.6%; Pred. No. 5.9e-47;
Matches 283; Conservative 157; Mismatches 375; Indels 249; Gaps 28;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         88 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 147

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
Db        148 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 198

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
Db        199 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 248

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
Db        249 GNVIMVRLLLDRGAQIETRTKDELTPLHCAARNGHLRISEILLDHGAPIQAKTKNGLSPI 308

Qy        216 HEAASKNRVEVCSLLLSYGA---DPTL---------LNC-HNKSA---IDLAPTPQLKER 259
Db        309 HMAAQGDHLDCVRLLLQYNAEIDDITLDHLTPLHVAAHCGHHRVAKVLLDKGAKPNSR-- 366

Qy        260 LAYEFKGHSLLQAAREADVTRI-----KKHLSLEMVNFKHPQTHETALHCAAASPYPKRK 314
Db        367 ---ALNGFTPLHIACKKNHIRVMELLLKMGASIDAVT----ESGLTPLHVASFMGHP--- 416

Qy        315 QICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRA 374
Db        417 PIVKSLLQREASPNVSNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCA 476

Qy        375 AYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAA 434
Db        477 ARIGHTNMVKLLLENNANPNLATTAGHTPLHI-------------------------AA 510

Qy        435 KAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGG 493
Db        511 REGHVETALALLEKEASQTCMTKKG--FTPLHVAAKYGKVRMAELLLEHDAHPNAAGKSG 568

Qy        494 LVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPT 553
Db        569 LTPLHVAVHHNHLDVVRLLLPRGGSPHSPALNGYTPLHIAAKQNQLEVARSLLQYGGSAN 628

Qy        554 KKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGR 612
Db        629 AESVQGVTPLHLAAQEGHAEMVALL-------------------LSKQANGNLGNKSG- 667

Qy        613 HSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 672
Db        668 -LTPLHLVAQEGHIPVADVLIKHGVTVDATTRMGYTPLHVASHYGNIKLVKFLLQHKADV 726

Qy        673 NATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSA 726
Db        727 NAKTKLGYSPLHQAAQQGHTDIVTLLLKHGASPNEVSSNGTTPLAIAKRLGYISVTDVLK 786

Qy        727 LLTAAMPPSALPSCYK---PQVLN-------------GVRSPGATA-------------- 756
Db        787 VVTDETTVQLITDKHRMSFPETVDEILDVSEDEGEDLGPRAERPDARDVDEEKELLDFVP 846

Qy        757 --DALSSGPSSP--------SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 806
```

```
Db        847 KLDQVVESPAIPRIPCVTPETVVIRADEPDQVSKEYDEDSLIPSSPATETSDNISPVASP 906

Qy        807 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVE- 865

Db        907 VHTGFLVSFM----------------------------VDARGGSMRGSRHNGLRVVVP 937

Qy        866 ------------RLISGQQGLNPYLTLNTSGSGTILIDLSP------------------- 894

Db        938 PRTCAAPTRITCRLVKPQKLPTPPPLAEEEGLASRVIALGPTGVQFLSPVIVEIPHPASQ 997

Qy        895 --DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCN 936

Db        998 GRGDRELVVLRSENGSVWKEHK---------NRYGESYLDQILN 1032


RESULT 11
O70511
ID   O70511      PRELIMINARY;      PRT;  2622 AA.
AC   O70511;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   270 kDa ANKYRIN G isoform (ANKYRING) (Fragment).
GN   ANK3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98417633; PubMed=9744885;
RA   Zhang X., Bennett V.;
RT   "Restriction of 480/270-kD ankyrin G to axon proximal segments
RT   requires multiple ankyrin G-specific domains.";
RL   J. Cell Biol. 142:1571-1581(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Carpenter S.S., Zhang X.;
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 934-1220 FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Skeletal muscle;
RX   MEDLINE=98330444; PubMed=9664041;
RA   Kordeli E., Ludosky M.A., Deprette C., Frappier T., Cartaud J.;
RT   "Ankyrin G is associated with the postsynaptic membrane and the
RT   sarcoplasmic reticulum in the skeletal muscle fibre.";
RL   J. Cell Sci. 111:2197-2207(1998).
DR   EMBL; AF102552; AAC78143.1; -.
DR   EMBL; AF065150; AAC18853.1; -.
DR   HSSP; P55273; 1BI8.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
FT   NON_TER    2622   2622
SQ   SEQUENCE   2622 AA;  284456 MW;  67B34830D3AC884E CRC64;

  Query Match             15.0%;  Score 864.5;  DB 11;  Length 2622;
  Best Local Similarity   30.5%;  Pred. No. 3.4e-46;
  Matches  276;  Conservative  123;  Mismatches  360;  Indels  147;  Gaps   23;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         85 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 144

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114

Db        145 YMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 204

Qy        115 LLESARSGNEEKMMALL----------TPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLL 164

Db        205 ALHIAARKDDTKAAALLLQNDTNADIESKMVVNRATESG--FTSLHIAAHYGNINVATLL 262

Qy        165 LQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRV 224

Db        263 LNRAAAVDFTARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHE 322

Qy        225 EVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKH 284

Db        323 QVVEMLLDRAA-PILSKTKN--------------------GLSPLHMATQGD------H 354

Qy        285 LSLEMVNFKH-------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTP 337

Db        355 LNCVQLLLQHNVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTP 411

Qy        338 LHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIS 397

Db        412 LHIACKKNRIRVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTN 471

Qy        398 LQGFTALQMG-----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-Q 449

Db        472 VRGETALHMAARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGA 531

Qy        450 SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVA 509

Db        532 SPNAATTSG--YTPLHLSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVA 589

Qy        510 ELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKD 568

Db        590 SLLLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKK 649

Qy        569 GDTDIQDLL------------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHST 615

Db        650 NQMDIATSLLEYGADANPVTRQGIASVHLAAQEGHVDMVSLLLSRNANVNLSNKSG--LT 707

Qy        616 PLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNAT 675

Db        708 PLHLAAQEDRVNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAK 767

Qy        676 DKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS-------ADDVSALL 728

Db        768 TKNGYTPLHQAAQQGHTHIINVLLQNNASPNELTVNGNTALAIARRLGYISVVDTLKVVT 827

Qy        729 TAAMPPSALPSCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLS-GSFSE 784

Db        828 EEIMTTTTITEKHKMNVPETMNEVLD-----------MSDDEVGKASAPEKLSDGEY-- 873

Qy        785 LSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHK 844

Db        874 -----ISDGEEG--------------------------------EDAITGDTDKYLGPQ 895

Qy        845 ELKEIG 850

Db        896 DLKELG 901


RESULT 12
Q7Z3L5
ID   Q7Z3L5      PRELIMINARY;      PRT;  1863 AA.
AC   Q7Z3L5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein DKFZp686H0688.
GN   DKFZP686H0688.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Human retina;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX537758; CAD97827.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   1863 AA;  204736 MW;  1F4C998E0F0A03DF CRC64;

  Query Match           15.0%;  Score 864;  DB 4;  Length 1863;
  Best Local Similarity  29.4%;  Pred. No. 2.2e-46;
  Matches  298;  Conservative  135;  Mismatches  412;  Indels  168;  Gaps   24;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :|: ||  |:|| :    |  || |  | |||| :|:: ||: ||:    :|||
Db         54 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPL 113

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL--- 115
              : || :  |||   ||::||  :   || |     ||     | |:|     | ::
Db        114 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLP 173

Qy        116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----------TPLHLAAGYNRVKIVQL 163
               |  |   :: |  |||     | | :| :                ||||:|| |   | :  |
Db        174 ALHIAARKDDTKSAALLLQ---NDHNADIQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 230

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
              ||  || |    :  : ||| |   |:  : :||:  |  ::|       |||| ||
Db        231 LLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 290

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 283
              :|  |||  || | |    |                    | | |  | : |     |
Db        291 DQVVELLLERGA-PLLARTKN--------------------GLSPLHMAAQGDHVECVK 328

Qy        284 HLSLEMVNFKHP-----QTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPL 338
              ||    :  | |       : |||| ||   :   :: :||| | || | :     |||
Db        329 HL----LQHKAPVDDVTLDYLTALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPL 381

Qy        339 HVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISL 398
              |:| :|    |:|::||: | : |:   | | :| ||: |||    |||  |  |:: ::
Db        382 HIACKKNRIKVMELLVKYGASIQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNI 441

Qy        399 QGFTALQMGN-----ENVQQLLQEGI----------------SLGNSEADRQLLE---- 432
              :| ||| |       | |: ||: |                || :|  : ||:
Db        442 RGETALHMAARAGQVEVVRCLLRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAH 501

Qy        433 --------------AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                           :|: | |:    |   : :   : :  |||| || |  : | :
Db        502 PDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKL 560

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
              |||  |   :  | || ||| |  | : :|| ||::  ||   :      :|||| || | :
Db        561 LLQRRAAADSAGKNGLTPLHVAAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQ 620

Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597
              :|   || :||:      : | ||| |  ::| ||:  ||    | :  : |  |
Db        621 MQIASTLLNYGAETNIVTKQGVTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----- 675

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                           | |||||  : : ||: | :|||| :|  | |  ||  |  ||
Db        676 ----------------TSLHLAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYG 718

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717
              :| :   |:|  | |||   |  :||||:|||:| |  :  :|| ||| |       | | |
Db        719 NVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALA 778

Qy        718 L------VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 771
              :      :|  |  ::|  : :     |  : :  |    : |           :|
Db        779 IAKRLGYISVVDTLKVVTEEVTTTTTTITEK----HKLNVPETMTEVL--------DVSD 826

Qy        772 ASSLDNLSGSFSEL----------SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RN 818
                 | ::|    |             |: ||  :| : |     |   |:  |    |:
Db        827 EEGDDTMTGDGGEYLRPEDLKELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRS 886

Qy        819 LGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYL 878
                |  |    : |:   :|  | :   ::  :    |  : ::|  | | |       |  |
Db        887 HTLSHAS--YLRDSAVMDDSVVIPSHQVSTLAKEAERNSYRLSWGTENLD------NVAL 938

Qy        879 TLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKI 931
              : :   ||      |  |: |      : | : : | |    |    |:| |:|
Db        939 SSSPIHSGRASPCLERDNSSF------LVSFMVDARGGAMRG---CRHNGLRI 982

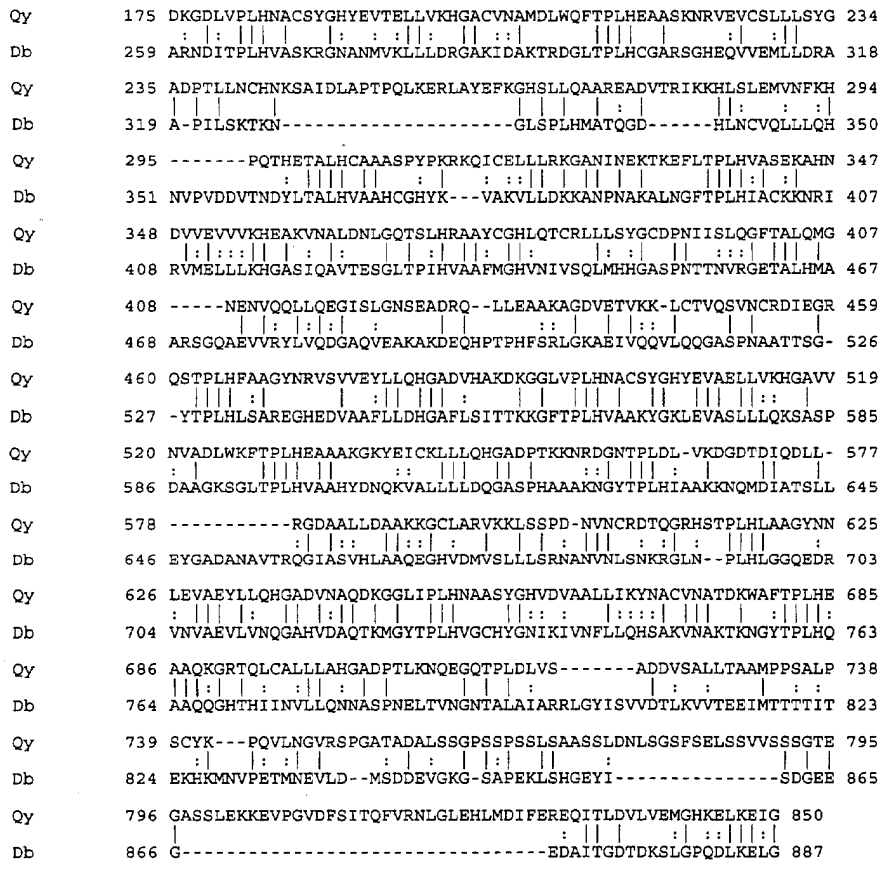

RESULT 13
O88521
ID   O88521      PRELIMINARY;      PRT;  1762 AA.
AC   O88521;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   190 kDa ankyrin isoform.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98395113; PubMed=9727010;
RA   Thevananther S., Kolli A.H., Devarajan P.;
RT   "Identification of a novel ankyrin isoform (AnkG190) in kidney and
RT   lung that associates with the plasma membrane and binds alpha-Na, K-
RT   ATPase.";
RL   J. Biol. Chem. 273:23952-23958(1998).
DR   EMBL; AF069525; AAC34809.1; -.
DR   HSSP; P55273; 1BI8.
DR   GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000169; SHprot_acsite.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 19.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1762 AA;  191899 MW;  9023280086A7BF4E CRC64;

  Query Match           15.0%;  Score 863.5;  DB 11;  Length 1762;
  Best Local Similarity  30.2%;  Pred. No. 2.1e-46;
  Matches  270;  Conservative  123;  Mismatches  369;  Indels  133;  Gaps   21;

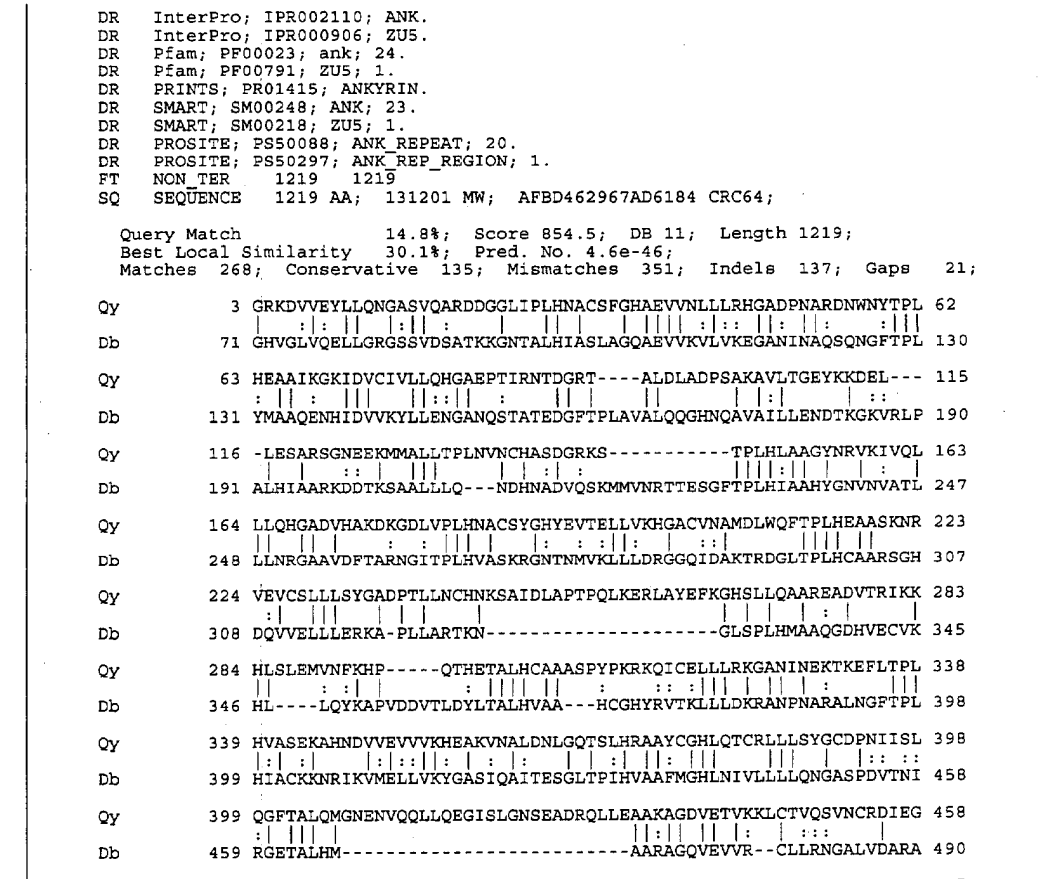
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:|      |   || |   | |||| :|: :||: ||:    :|||
Db         79 GHVEVVSELLQREANVDQPTKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 138

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|    || :||  ::    || | | :|      |     :|  : |  :
Db        139 YMAAQENHLEVVRFLLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 198

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |     |  :: |  |||    : |           ||||:|| |  : :  |||    | |
Db        199 ALHIAARKDDTKAAALLLQNDTNADIESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 258
```



RESULT 14
Q8C8R3
ID Q8C8R3 PRELIMINARY; PRT; 1219 AA.
AC Q8C8R3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ankyrin 2 (Fragment).
GN ANK2 OR AI835472.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK044634; BAC32012.1; -.
DR MGD; MGI:88025; Ank2.



DR InterPro; IPR002110; ANK.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 23.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 20.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
FT NON_TER 1219 1219
SQ SEQUENCE 1219 AA; 131201 MW; AFBD462967AD6184 CRC64;

Query Match 14.8%; Score 854.5; DB 11; Length 1219;
Best Local Similarity 30.1%; Pred. No. 4.6e-46;
Matches 268; Conservative 135; Mismatches 351; Indels 137; Gaps 21;

```
Db        826 KHKLNVPETMTE---------VLDVSDEEGDDTVTGDGGEYLRPEDLKELG 867


RESULT 15
Q13768
ID   Q13768      PRELIMINARY;      PRT;  1719 AA.
AC   Q13768;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alt. ankyrin (Variant 2.2).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hematopoietic;
RX   MEDLINE=90158830; PubMed=2137557;
RA   Lux S.E., John K.M., Bennett V.;
RT   "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
RT   structure with homology to tissue-differentiation and cell-cycle
RT   control proteins.";
RL   Nature 344:36-42(1990).
DR   EMBL; X16609; CAA34611.1; -.
DR   HSSP; Q00420; 1AWC.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1719 AA;  189011 MW;  F63465D16D975CBF CRC64;

  Query Match             14.6%;  Score 842;  DB 4;  Length 1719;
  Best Local Similarity   25.5%;  Pred. No. 5e-45;
  Matches  279;  Conservative  152;  Mismatches  356;  Indels  306;  Gaps   27;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         89 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 148

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
Db        149 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 199

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS----------------TPLHLAAGY 155
Db        200 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 249

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
Db        250 GNVIMVRLLLDRGAQIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
Db        310 HMAAQGDHLDCVRLLLQYDAE----------ID------------------------- 332

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
Db        333 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 372

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395

Db        373 TPLHIACKKNHVRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 432

Qy        396 ISLQGFTALQMG-----NENVQQLLQE----------------GISLGNSEADRQLLE- 432
Db        433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLEN 492

Qy        433 ----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db        493 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVR 550

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
Db        551 VAELLLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 609

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
Db        610 AKQNQVEVARSLLQYGGSANAESVQGVTPLHLAAQEGHAEMVALL---------------- 654

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
Db        655 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHV 707

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
Db        708 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDG 767

Qy        713 QTPLDL--------------VSADDVSALLTAAMPPSALPSC------------------ 740
Db        768 TTPLAIAKRLGYISVTDVLKVVTDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISF 827

Qy        741 ------------------YKPQVLNGVRSPGATADALSSGPSSPSSLS-----AASSLDN 777
Db        828 KAERRDSRDVDEEKELLDFVPKLDQVVESPAI--------PRIPCAMPETVVIRSEEQEQ 879

Qy        778 LSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDV 837
Db        880 ASKEYDEDSLIPSSPATETSDNISPVASPVHTGFLVSFM---------------------- 918

Qy        838 LVEMGHKELKEIGINAYGHRHKLIKGVE------------RLISGQQGLNPYLTLNTSG 884
Db        919 --------VDARGGSMRGSRHNGLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEG 970

Qy        885 SGTILIDLSP--------------------DDKEFQSVEEEMQSTVREHRDGGHAGGIF 923
Db        971 LASRIIALGPTGAQFLSPVIVEIPHFASHGRGDRELVVLRSENGSVWKEHR--------- 1021

Qy        924 NRYNILKIQKVCN 936
Db       1022 SRYGESYLDQILN 1034


RESULT 16
P97582
ID   P97582      PRELIMINARY;      PRT;  843 AA.
AC   P97582;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ankyrin (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Brain;
RA   Cosentino M.T., Jones O.T.;
RT   "Rat brain ankyrin membrane binding domain.";
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U65916; AAB47551.1; -.
DR   HSSP; Q00420; 1AWC.
DR   InterPro; IPR002110; ANK.
```

```
DR   Pfam; PF00023; ank; 23.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER       1      1
FT   NON_TER     843    843
SQ   SEQUENCE   843 AA;  89982 MW;  DC5A6AC78825D223 CRC64;

  Query Match             14.6%;  Score 841.5;  DB 11;  Length 843;
  Best Local Similarity   29.7%;  Pred. No. 1.8e-45;
  Matches  265;  Conservative  138;  Mismatches  351;  Indels  137;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |   :|: ||  |:|| :    |   || |   | |||| :|:: ||: ||:    :|||
Db         40 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPL 99

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL--- 115
              : || :  |||   ||::||  :    || |    ||       | |:|      | ::
Db        100 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLP 159

Qy        116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS----------TPLHLAAGYNRVKIVQL 163
               |  |   ::  |  |||     | | :| :              ||||:|| |  | : |
Db        160 ALHIAARKDDTKSAALLLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 216

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
              ||  || |    :  : ||| |   |:  : :||:  |  ::|      |||| ||
Db        217 LLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 276

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 283
              :|  |||  || | | |     |                  | | |  | : |       |
Db        277 DQVVELLLERGA-PLLARTKN--------------------GLSPLHMAAQGDHVECVK 314

Qy        284 HLSLEMVNFKHP-----QTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPL 338
              ||    :  | |       : |||| ||   :    :: :||| | || | :     |||
Db        315 HL----LQHKAPVDDVTLDYLTALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPL 367

Qy        339 HVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISL 398
              |:| :|    |:|::||: | : |:   | | :  ||: |||    |||  |  |:: ::
Db        368 HIACKKNRIKVMELLVKYGAYIQAITESGLTPIPVAAFMGHLNIVLLLLQNGASPDVTNI 427

Qy        399 QGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEG 458
              :| ||| |                          ||:||:|| |:  | :::    |
Db        428 RGETALHM-------------------------AARAGEVEVVR--CLLRNGALVDARA 459

Qy        459 R-QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517
              | : |||| |:   :  :|: |||| |   |    | |||  :   |  :|| :|:: ||
Db        460 REEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTNGYTPLHISAREGQVDVASVLLEAGA 519

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLL 577
              ::|    ||||| ||  |  :: |||||  |      ::| ||| :    |
Db        520 AHSLATKKGFTPLHVAAKYGSLDVAKLLLQRRAAADSAGKNGLTPLHVAAHYDN------ 573

Qy        578 RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGA 637
                    :|  |  ::| :||       |     ||||:||  | :::|  || :||
Db        574 ----------QKVALLLLEKGASP-----HATAKNGYTPLHIAAKKNQMQIASTLLNYGA 618

Qy        638 DVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCAL 697
              : |    | |: ||| |:  || |:  |::  | :: : |    | || ||:: :  :  :
Db        619 ETNTVTKQGVTPLHLASQEGHTDMVTLVLEKGANIHMSTKSGLTSLHLAAEEDKVNVADI 678

Qy        698 LLAHGADPTLKNQEGQTPLDLV----SADDVSALLTAAMPPSA-LPSCYKP--------- 743
              |  ||||     : |||| :      :    |: ||       :|   : | |
Db        679 LTKHGADQDAYTKLGYTPLIVACHYGNVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGH 738

Qy        744 -QVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEK 802
               ::| :   ||  :| ::  :  ::|: |   |  :  |  :    ||:    |   ::: :
Db        739 THIINVLLQHGAKPNATTA--NGNTALAIAKRLGYI--SVVDTLKVVTEEVTTTTTTITE 794

Qy        803 KEVPGVDFSITQFVRNLGLEHLMDIFERE---QITLDVLVEMGHKELKEIG 850
              |       ::|:          ::|: : |     :| |    :  ::|||:|
Db        795 KHKLNAPETMTE---------VLDVSDEEGDDTVTGDGGEYLRPEDLKELG 836
```

```
RESULT 17
Q99407
ID   Q99407      PRELIMINARY;      PRT;  1856 AA.
AC   Q99407;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ankyrin.
GN   ANK1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97382244; PubMed=9235914;
RA   Gallagher P.G., Tse W.T., Scarpa A.L., Lux S.E., Forget B.G.;
RT   "Structure and organization of the human ankyrin-1 gene. Basis for
RT   complexity of pre-mRNA processing.";
RL   J. Biol. Chem. 272:19220-19228(1997).
DR   EMBL; U50133; AAB47805.1; -.
DR   EMBL; U50092; AAB47805.1; JOINED.
DR   EMBL; U50093; AAB47805.1; JOINED.
DR   EMBL; U50094; AAB47805.1; JOINED.
DR   EMBL; U50095; AAB47805.1; JOINED.
DR   EMBL; U50096; AAB47805.1; JOINED.
DR   EMBL; U50097; AAB47805.1; JOINED.
DR   EMBL; U50098; AAB47805.1; JOINED.
DR   EMBL; U50099; AAB47805.1; JOINED.
DR   EMBL; U50100; AAB47805.1; JOINED.
DR   EMBL; U50101; AAB47805.1; JOINED.
DR   EMBL; U50102; AAB47805.1; JOINED.
DR   EMBL; U50103; AAB47805.1; JOINED.
DR   EMBL; U50104; AAB47805.1; JOINED.
DR   EMBL; U50105; AAB47805.1; JOINED.
DR   EMBL; U50106; AAB47805.1; JOINED.
DR   EMBL; U50107; AAB47805.1; JOINED.
DR   EMBL; U50108; AAB47805.1; JOINED.
DR   EMBL; U50109; AAB47805.1; JOINED.
DR   EMBL; U50110; AAB47805.1; JOINED.
DR   EMBL; U50111; AAB47805.1; JOINED.
DR   EMBL; U50112; AAB47805.1; JOINED.
DR   EMBL; U50113; AAB47805.1; JOINED.
DR   EMBL; U50114; AAB47805.1; JOINED.
DR   EMBL; U50115; AAB47805.1; JOINED.
DR   EMBL; U50116; AAB47805.1; JOINED.
DR   EMBL; U50117; AAB47805.1; JOINED.
DR   EMBL; U50118; AAB47805.1; JOINED.
DR   EMBL; U50119; AAB47805.1; JOINED.
DR   EMBL; U50120; AAB47805.1; JOINED.
DR   EMBL; U50121; AAB47805.1; JOINED.
DR   EMBL; U50122; AAB47805.1; JOINED.
DR   EMBL; U50123; AAB47805.1; JOINED.
DR   EMBL; U50124; AAB47805.1; JOINED.
DR   EMBL; U50125; AAB47805.1; JOINED.
DR   EMBL; U50126; AAB47805.1; JOINED.
DR   EMBL; U50127; AAB47805.1; JOINED.
DR   EMBL; U50128; AAB47805.1; JOINED.
DR   EMBL; U50129; AAB47805.1; JOINED.
DR   EMBL; U50130; AAB47805.1; JOINED.
DR   EMBL; U50131; AAB47805.1; JOINED.
DR   EMBL; U50132; AAB47805.1; JOINED.
DR   HSSP; Q00420; 1AWC.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
```

```
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1856 AA;  203466 MW;  4A4E3A794DD6401F CRC64;

  Query Match             14.6%;  Score 840.5;  DB 4;  Length 1856;
  Best Local Similarity   29.9%;  Pred. No. 7e-45;
  Matches  240;  Conservative  122;  Mismatches  264;  Indels  177;  Gaps   20;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |      :|||
Db         56 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 115

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                | :|  :|   |: :|: :  :          || | |||          |||
Db        116 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 166

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
              ::|  :::|          |||::  |  :   :                    ||||:|:
Db        167 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 216

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
              | :|:||| || :  | | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        217 GNVIMVRLLLDRGAQIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 276

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:            ||
Db        277 HMAAQGDHLDCVRLLLQYDAE-----------ID-------------------------- 299

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                  | | || ||   :    :: ::|| |||  | :
Db        300 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 339

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : || |  ||:
Db        340 TPLHIACKKNHVRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 399

Qy        396 ISLQGFTALQMG-----NENVQQLLQE----------------GISLGNSEADRQLLE- 432
              :::  | | |       |  : |||                    :|::   : |||
Db        400 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLEN 459

Qy        433 ----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
                           ||: | ||||  |   : |  |   :|  ||||  || | :|
Db        460 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVR 517

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
              | | ||:  |  :|  | || |||  | : : :: :||:  |   : :  |  :|||| |
Db        518 VAELLLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 576

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
              | : : |: : |||:|     ::  | ||| |  ::|  ::  ||
Db        577 AKQNQVEVARSLLQYGGSANAESVQGVTPLHLAAQEGHAEMVALL--------------- 621

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
                  ||   | |   :  |   ||||| |   :: ||: |::||  |:|  :  |  |||
Db        622 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHV 674

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
              |: ||:: :   |::: | |||  |  ::|||:|||:| | :  ||| :|| |   : :|
Db        675 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDG 734

Qy        713 QTPLDL------VSADDVSALLT 729
              |||  :      :|  ||  ::|
Db        735 TTPLAIAKRLGYISVTDVLKVVT 757


RESULT 18
Q61302
ID   Q61302      PRELIMINARY;      PRT;  1848 AA.
AC   Q61302;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Erythroid ankyrin.
GN   ANK1 OR ANK-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Nervous system;
RX   MEDLINE=93252825; PubMed=8486643;
RA   Birkenmeier C.S., White R.A., Peters L.L., Hall E.J., Lux S.E.,
RA   Barker J.E.;
RT   "Complex patterns of sequence variation and multiple 5' and 3' ends
RT   are found among transcripts of the erythroid ankyrin gene.";
RL   J. Biol. Chem. 268:9533-9540(1993).
DR   EMBL; X69063; CAA48801.1; -.
DR   PIR; S37771; S37771.
DR   HSSP; Q00420; 1AWC.
DR   MGD; MGI:88024; Ank1.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1848 AA;  202577 MW;  5744BECBF9EBA056 CRC64;

  Query Match             14.6%;  Score 839.5;  DB 11;  Length 1848;
  Best Local Similarity   29.5%;  Pred. No. 8.1e-45;
  Matches  237;  Conservative  127;  Mismatches  262;  Indels  177;  Gaps   20;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |      :|||
Db         93 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 152

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                | :|  :|   |: :|: :  :          || | |||          |||
Db        153 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 203

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
              ::|  :::|          |||::  |  :   :                    ||||:|:
Db        204 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 253

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
              | :|:||| || :  | | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        254 GNVIMVRLLLDRGAQIETRTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 313

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:            ||
Db        314 HMAAQGDHLDCVRLLLQYNAE-----------ID-------------------------- 336

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                  | | || ||   :    :: ::|| |||  | :
Db        337 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 376

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : || |  ||:
Db        377 TPLHIACKKNHIRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 436

Qy        396 ISLQGFTALQM----GNENVQQLLQEGISLGNSEAD----------------RQLLE- 432
```

```
Db        437 SNVKVETPLHMAARAGHTEVAKYLLQNKAKANAKAKDDQTPLHCAARIGHTGMVKLLLEN 496

Qy        433 -----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474

Db        497 GASPNLATTAGHTPLHTAAREGHVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVR 554

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533

Db        555 LAELLLEHDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 613

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592

Db        614 AKQNQIEVARSLLQYGGSANAESVQGVTPLHLAAQEGHTEMVALL---------------- 658

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652

Db        659 -----LSKQANGNLGNKSG--LTPLHLVSQEGHVPVADVLIKHGVTVDATTRMGYTPLHV 711

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712

Db        712 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNG 771

Qy        713 QTPLDL------VSADDVSALLT 729

Db        772 TTPLAIAKRLGYISVTDVLKVVT 794
```

```
RESULT 19
Q9V4B1
ID   Q9V4B1      PRELIMINARY;      PRT;  1549 AA.
AC   Q9V4B1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ank protein.
GN   ANK OR CG1651 OR CG32008.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003845; AAF59369.2; -.
DR   HSSP; P80144; 2MYO.
DR   FlyBase; FBgn0011747; Ank.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0045170; C:spectrosome; IDA.
DR   GO; GO:0007016; P:cytoskeletal anchoring; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1549 AA;  170167 MW;  3ECEC852B1484DFF CRC64;
```

```
  Query Match           14.5%;  Score 836;  DB 5;  Length 1549;
  Best Local Similarity  27.4%;  Pred. No. 1e-44;
  Matches  298;  Conservative  166;  Mismatches  469;  Indels  154;  Gaps   32;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         84 GYVDICCELLRRGIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPL 143

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114

Db        144 YMAAQENHDNCCRTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLENDVRGKVRLP 203

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174

Db        204 ALHIAAKKNDVNAAKLLLQHDPNADIVSKSGFTPLHIAAHYGNVDIATLLLNNKADVNYV 263

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234

Db        264 AKHNITPLHVACKWGKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLLQQN 323

Qy        235 ADPTLLNCHN-KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE----- 288

Db        324 A-PILTKTKNGLSALHMAAQGE------HDEAAHLLLDNKAPVDEVTVDYLTALHVAAHC 376

Qy        289 --------MVNFK-HPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335

Db        377 GHVKVAKLLLDYKANPNARALNGFTPLHIACKK---NRIKMVELLIKHGANIGATTESGL 433

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395

Db        434 TPLHVASFMGCINIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDA 492

Qy        396 ISLQGFTAL----QMGNEN-VQQLLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTV 448

Db        493 IAREGQTPLHVASRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENIVQVLLEN 552

Qy        449 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 508

Db        553 GAEN-NAVTKKGFTPLHLACKYGKQNVVQILLQNGASIDFQGKNDVTPLHVATHYNNPSI 611
```

```
Qy        509 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VK 567
Db        612 VELLLKNGSSPNLCARNGQCAIHIACKKNYLEIAMQLLQHGADVNIISKSGFSPLHLAAQ 671

Qy        568 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLE 627
Db        672 GGNVDMVQLLLEYGVISAAAKNGL----------------------TPLHVAAQEGHVL 708

Qy        628 VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA 687
Db        709 VSQILLEHGANISERTRNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAA 768

Qy        688 QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLN 747
Db        769 QQGHIMIINLLLRHKANPNALTKDGNTALHIASN------LGYVTVMESLKIVTSTSVIN 822

Qy        748 G----------VRSPGATADALSSGPSSPSSLSAAS-------SLDNLSGSFSELSSVVS 790
Db        823 SNIGAIEEKLKVMTPELMQETLLSDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFD 882

Qy        791 SSGTE----GASSLEKKEVPGVDFSITQFVRNLGLEHLMD--IFEREQITLDVLVEMGHK 844
Db        883 TTNTDHDLTDVSVLNKKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF--- 936

Qy        845 ELKEIGINAYGHRHKLIKGVE------------RLISGQQGLNPYLTLNTSGSGTILID 891
Db        937 LVDARGGSMRGYRHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSRILE 996

Qy        892 LSPDDKEFQS---VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 948
Db        997 MSPVDGMFLSPITLEVPHYGTLRK----------NEREII-ILRSDNGESWREH-NLYK 1043

Qy        949 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 1008
Db       1044 DIIGEDINQTEE---FHSDRIVR-IVTQNVPHFFAVV-------------SRVRQEVHV 1085

Qy       1009 IG--GGT 1013
Db       1086 IGPDGGT 1092
```

```
RESULT 20
Q24241
ID   Q24241      PRELIMINARY;      PRT;  1549 AA.
AC   Q24241;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ankyrin.
GN   ANK OR CG1651 OR CG32008.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A631;
RX   MEDLINE=95024098; PubMed=7937942;
RA   Dubreuil R.R., Yu J.;
RT   "Ankyrin and beta-spectrin accumulate independently of alpha-spectrin
RT   in Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:10285-10289(1994).
DR   EMBL; L35601; AAC37208.1; -.
DR   PIR; T13940; T13940.
DR   HSSP; P80144; 2MYO.
DR   FlyBase; FBgn0011747; Ank.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0045170; C:spectrosome; IDA.
DR   GO; GO:0007016; P:cytoskeletal anchoring; IDA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 23.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1549 AA;  170096 MW;  587FB6BF88E657CE CRC64;
```

```
  Query Match           14.5%;  Score 834;  DB 5;  Length 1549;
  Best Local Similarity   27.4%;  Pred. No. 1.4e-44;
  Matches  298;  Conservative  165;  Mismatches  470;  Indels  154;  Gaps   32;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         84 GYVDICCELLRRGIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPL 143

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
Db        144 YMAAQENHDNCCRTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLENDVRGKVRLP 203

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
Db        204 ALHIAAKKNDVNAAKLLLQHDPNADIVSKSGFTPLHIAAHYGNVDIATLLLNNKADVNYV 263

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
Db        264 AKHNITPLHVACKWGKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLLQQN 323

Qy        235 ADPTLLNCHN-KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE----- 288
Db        324 A-PILTKTKNGLSALHMAAQGE------HDEAAHLLLDNKAPVDEVTVDYLTALHVAAHC 376

Qy        289 --------MVNFK-HPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
Db        377 GHVKVAKLLLDYKANPNARALNGFTPLHIACKK---NRIKMVELLIKHGANIGATTESGL 433

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
Db        434 TPLHVASFMGCINIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDA 492

Qy        396 ISLQGFTAL----QMGNEN-VQQLLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTV 448
Db        493 IVREGQTPLHVASRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENIVQVLLEN 552

Qy        449 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 508
Db        553 GAEN-NAVTKKGFTPLHLACKYGKQNVVQILLQNGASIDFQGKNDVTPLHVATHYNNPSI 611

Qy        509 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VK 567
Db        612 VELLLKNGSSPNLCARNGQCAIHIACKKNYLEIAMQLLQHGADVNIISKSGFSPLHLAAQ 671

Qy        568 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLE 627
Db        672 GGNVDMVQLLLEYGVISAAAKNGL----------------------TPLHVAAQEGHVL 708

Qy        628 VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA 687
Db        709 VSQILLEHGANISERTRNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAA 768

Qy        688 QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLN 747
Db        769 QQGHIMIINLLLRHKANPNALTKDGNTALHIASN------LGYVTVMESLKIVTSTSVIN 822

Qy        748 G----------VRSPGATADALSSGPSSPSSLSAAS-------SLDNLSGSFSELSSVVS 790
Db        823 SNIGAIEEKLKVMTPELMQETLLSDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFD 882
```

```
Qy        791 SSGTE----GASSLEKKEVPGVDFSITQFVRNLGLEHLMD--IFEREQITLDVLVEMGHK 844
              :: |:       | | |||:   : |  :      : |  |  :  | |: |  ||
Db        883 TTNTDHDLTDVSVLNKKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF--- 936

Qy        845 ELKEIGINAYGHRHKLIKGVE------------RLISGQQGLNPYLTLNTSGSGTILID 891
               :   | :  |:||  :: :              | :  |: :||   :      : :::
Db        937 LVDARGGSMRGYRHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSRILE 996

Qy        892 LSPDDKEFQS---VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 948
              :|| |  | |   :|      |:|:          |   |: | :  | : |  : :  |
Db        997 MSPVDGMFLSPITLEVPHYGTLRK----------NEREII-ILRSDNGESWREH-NLYK 1043

Qy        949 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 1008
              ::  |: |   |   ||     |  |: :     | :                |: | |:
Db       1044 DIIGEDINQTEE---FHSDRIVR-IVTQNVPHFFAVV-------------SRVRQEVHV 1085

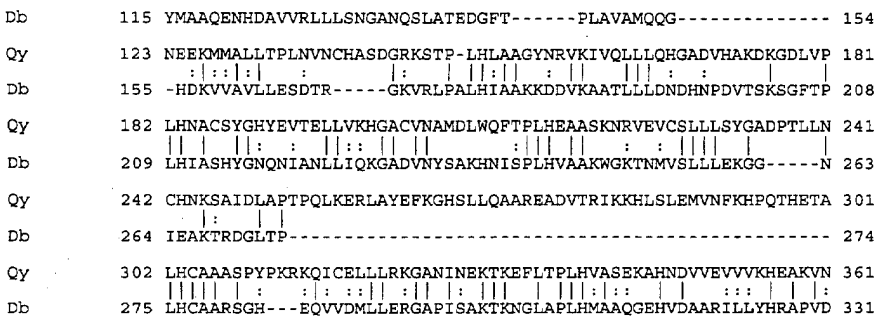
Qy       1009 IG--GGT 1013
              ||  |||
Db       1086 IGPDGGT 1092
```

```
RESULT 21
Q9NCP8
ID   Q9NCP8      PRELIMINARY;      PRT;  1159 AA.
AC   Q9NCP8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ankyrin 2 (CG7462-PC).
GN   ANK2 OR CG7462.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20304926; PubMed=10844021;
RA   Bouley M., Tian M.-Z., Paisley K., Shen Y.-C., Malhotra J.D.,
RA   Hortsch M.;
RT   "The L1-type cell adhesion molecule neuroglian influences the
RT   stability of neural ankyrin in the Drosophila embryo but not its
RT   axonal localization.";
RL   J. Neurosci. 20:4515-4523(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF190635; AAF73309.1; -.
DR   EMBL; AE003558; AAN12046.1; -.
DR   HSSP; P42773; 1IHB.
DR   FlyBase; FBgn0017645; Ank2.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   PROSITE; PS50088; ANK_REPEAT; 21.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1159 AA;  125769 MW;  9ED146C9E483DE58 CRC64;

  Query Match             14.2%;  Score 820;  DB 5;  Length 1159;
  Best Local Similarity   27.9%;  Pred. No. 7.1e-44;
  Matches  268;  Conservative  138;  Mismatches  373;  Indels  180;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |   ||  ||: || | :    |   || |   | ||| ||| | |  | :    :|||
Db         55 GHIHVVSELLRRGAIVDSATKKGNTALHIASLAGQEEVVKLLLEHNASVNVQSQNGFTPL 114

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
                : || :    |  :|| :||  ::   || |      | | |:  |
```



```
Db        115 YMAAQENHDAVVRLLLSNGANQSLATEDGFT------PLAVAMQQG-------------- 154

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181

Db        155 -HDKVVAVLLESDTR-----GKVRLPALHIAAKKDDVKAATLLLDNDHNPDVTSKSGFTP 208

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241

Db        209 LHIASHYGNQNIANLLIQKGADVNYSAKHNISPLHVAAKWGKTNMVSLLLEKGG-----N 263

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301

Db        264 IEAKTRDGLTP---------------------------------------------------- 274

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361

Db        275 LHCAARSGH---EQVVDMLLERGAPISAKTKNGLAPLHMAAQGEHVDAARILLYHRAPVD 331

Qy        362 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN-----VQQLLQ 416

Db        332 EVTVDYLTALHVAAHCGHVRVAKLLLDRNADANARALNGFTPLHIACKKNRLKVVELLLR 391

Qy        417 EGIS-----------------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR 459

Db        392 HGASISATTESGLTPLHVAAFMGCMNIVIYLLQHDASPDVPTV----------------R 435

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519

Db        436 GETPLHLAARANQTDIIRILLRNGAQVDARAREQQTPLHIASRLGNVDIVMLLLQHGAQV 495

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLR 578

Db        496 DATTKDMYTALHIAAKEGQDEVAAVLIENGAALDAATKKGFTPLHLTAKYGHIKVAQLLL 555

Qy        579 GDAALLDAAKKG----------------CLARVKKLSSPDNVNCRDTQGRHSTPLHLAAG 622

Db        556 QKEADVDAQGKNGVTPLHVACHYNNQQVALLLLEKGASP-----HATAKNGHTPLHIAAR 610

Qy        623 YNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTP 682

Db        611 KNQMDIATTLLEYGALANAESKAGFTPLHLSSQEGHAEISNLLIEHKAAVNHPAKNGLTP 670

Qy        683 LHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS----ADDVSALL-TAAMPPSAL 737

Db        671 MHLCAQEDNVNVAEILEKNGANIDMATKAGYTPLHVASHFGQANMVRFLLQNGANVDAAT 730

Qy        738 PSCYKP----------QVLNGVRSPGATADALSSGPSSP----SSLSAASSLDNLSGSFS 783

Db        731 SIGYTPLHQTAQQGHCHIVNLLLEHKANANAQTVNGQTPLHIARKLGYISVLDSLKTITK 790

Qy        784 ELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNL----GLEHLMDIFEREQITLDVLV 839

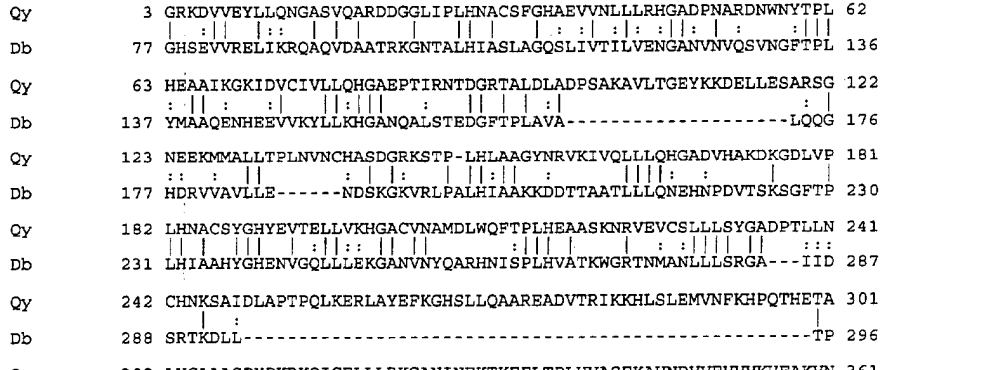
Db        791 EDETAAAPSQAE-----EKYRVVAPEAMHESFMSDSEEEGGEDNMLSDQPYRYLTVDEMK 845

Qy        840 EMGHKEL-------KEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILID 891

Db        846 SLGDDSLPIDVTRDERMDSNRMTQSAEYASGVPPTI-GEEVISPHKTQVYGSSPKATVD 903
```

RESULT 22
Q17487
ID Q17487 PRELIMINARY; PRT; 1809 AA.
AC Q17487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE E. elegans ankyrin-related unc-44 (GB:U21734).
GN B0350.2 OR UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gattung S.;
RT "The sequence of C. elegans cosmid B0350.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U50071; AAA93446.1; -.
DR HSSP; P42773; 1IHB.
DR WormPep; B0350.2b; CE06703.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001360; Glyco_hydro_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW ANK repeat; Repeat.
SQ SEQUENCE 1809 AA; 198193 MW; AD86349AA64534F8 CRC64;

Query Match 14.1%; Score 814.5; DB 5; Length 1809;
Best Local Similarity 26.1%; Pred. No. 3.2e-43;
Matches 301; Conservative 156; Mismatches 418; Indels 279; Gaps 36;



```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122

Db        137 YMAAQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQG 176

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181

Db        177 HDRVVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTP 230

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241

Db        231 LHIAAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IID 287

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301

Db        288 SRTKDLL---------------------------------------------------TP 296

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
```

```
Db        297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy        362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410

Db        354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411

Qy        411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457

Db        412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517

Db        457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576

Db        516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy        577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633

Db        576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy        634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693

Db        611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670

Qy        694 LCALLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPS 739

Db        671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy        740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796

Db        730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy        797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852

Db        779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

Qy        853 ----------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG 884

Db        836 ETMFSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGG 895

Qy        885 ------SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQ 932

Db        896 AIPKINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLH 950

Qy        933 KVCNKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DE 980

Db        951 GAGPEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDA 1010

Qy        981 RHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKS 1040

Db       1011 RGGAMRGCRHSGVRIIVPPRKASQ----------PIR-----------VTCRY-LRKD 1046

Qy       1041 FLQFSAMKMAHSPP 1054

Db       1047 -------KLAHPPP 1053


RESULT 23
Q8MQG0
ID   Q8MQG0      PRELIMINARY;      PRT;  1841 AA.
AC   Q8MQG0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   UNC-44 protein (corresponding sequence B0350.2e).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Gattung S.;
RT   "The sequence of C. elegans cosmid B0350.";
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAM75382.1; -.
DR   WormPep; B0350.2e; CE31147.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 23.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1841 AA;  201234 MW;  24CC6020DDCE0FC8 CRC64;

  Query Match             14.1%;  Score 814.5;  DB 5;  Length 1841;
  Best Local Similarity   26.1%;  Pred. No. 3.3e-43;
  Matches  301;  Conservative  156;  Mismatches  418;  Indels  279;  Gaps   36;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122

Db        137 YMAAQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQG 176

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181

Db        177 HDRVVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTP 230

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241

Db        231 LHIAAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IID 287

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301

Db        288 SRTKDLL-----------------------------------------------------TP 296

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361

Db        297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy        362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410

Db        354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411
```

```
Qy        411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457
Db        412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517
Db        457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576
Db        516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy        577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633
Db        576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy        634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693
Db        611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670

Qy        694 LCALLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPS 739
Db        671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy        740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
Db        730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy        797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852
Db        779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

Qy        853 ----------AYGHRHKLI------------------KGVERLISGQQGLNPYLTLNTSG 884
Db        836 ETMFSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGG 895

Qy        885 ------SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQ 932
Db        896 AIPKINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLH 950

Qy        933 KVCNKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DE 980
Db        951 GAGPEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDA 1010

Qy        981 RHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKS 1040
Db       1011 RGGAMRGCRHSGVRIIVPPRKASQ----------PIR------------VTCRY-LRKD 1046

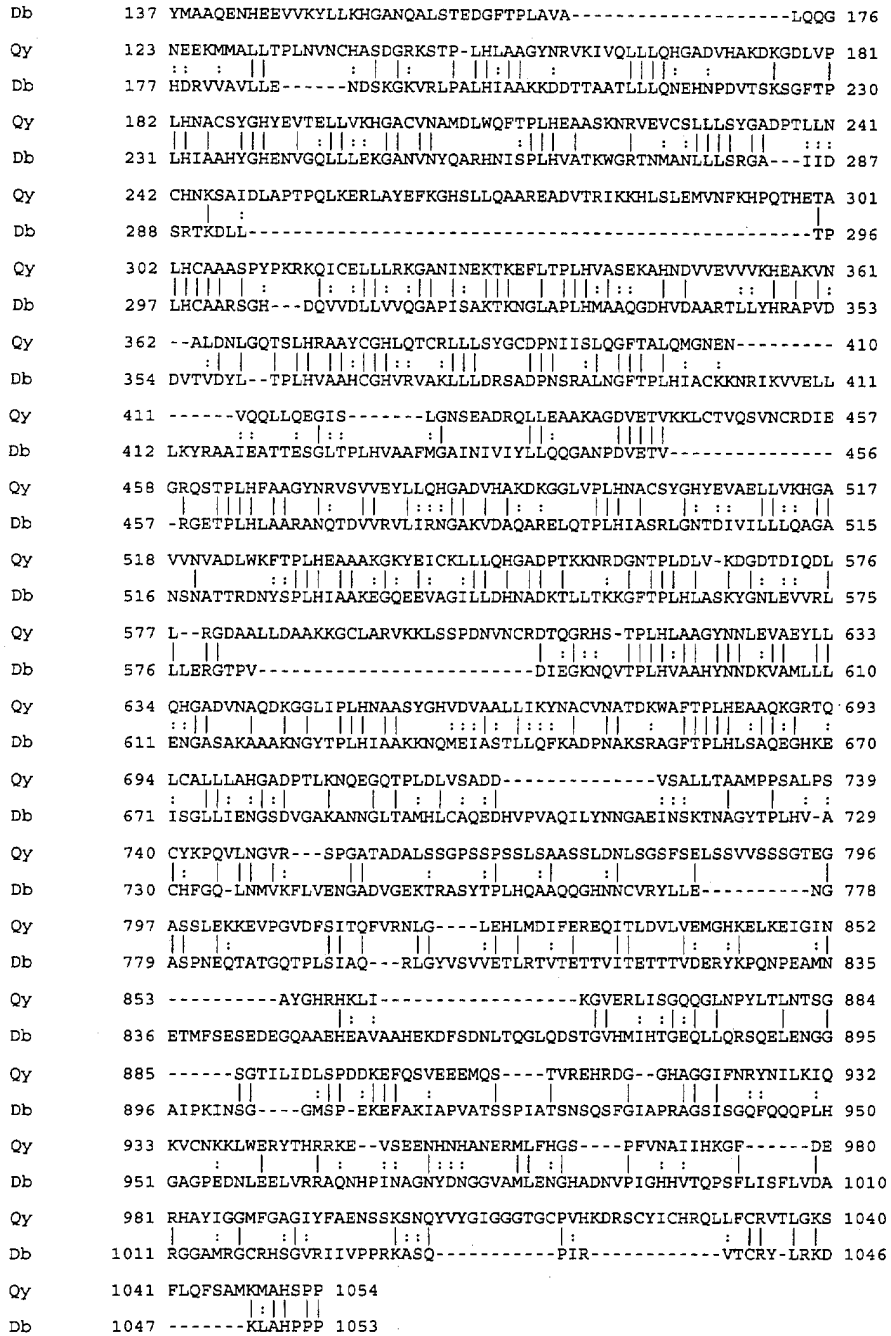
Qy       1041 FLQFSAMKMAHSPP 1054
Db       1047 -------KLAHPPP 1053
```

RESULT 24
Q17486
ID Q17486 PRELIMINARY; PRT; 1867 AA.
AC Q17486; O02516;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C. ELEGANS ankyrin-related UNC-44 (GB:U21734).
GN B0350.2 OR UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gattung S.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RX MEDLINE=95263663; PubMed=7744957;
RA Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA Sobery A.;
RT "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT guidance in Caenorhabditis elegans.";
RL J. Cell Biol. 129:1081-1092(1995).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Otsuka A.J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Otsuka A.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U50071; AAA93444.1; -.
DR EMBL; U39847; AAB41828.1; -.
DR HSSP; P42773; 1IHB.
DR WormPep; B0350.2c; CE06704.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001360; Glyco_hydro_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW ANK repeat; Repeat.
SQ SEQUENCE 1867 AA; 204253 MW; 4689A2104623B4C6 CRC64;

Query Match 14.1%; Score 814.5; DB 5; Length 1867;
Best Local Similarity 26.1%; Pred. No. 3.4e-43;
Matches 301; Conservative 156; Mismatches 418; Indels 279; Gaps 36;

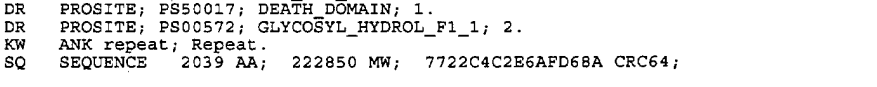
```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
```



RESULT 25
Q17489
ID Q17489 PRELIMINARY; PRT; 2039 AA.
AC Q17489;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C. elegans ankyrin-related unc-44 (GB:U21734).
GN B0350.2 OR UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gattung S.;
RT "The sequence of C. elegans cosmid B0350.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U50071; AAA93443.1; -.
DR PIR; T15347; T15347.
DR HSSP; P42773; 1IHB.
DR WormPep; B0350.2a; CE06702.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001360; Glyco_hydro_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 21.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW ANK repeat; Repeat.
SQ SEQUENCE 2039 AA; 222850 MW; 7722C4C2B6AFD68A CRC64;

Query Match 14.1%; Score 814.5; DB 5; Length 2039;
Best Local Similarity 26.1%; Pred. No. 3.9e-43;
Matches 301; Conservative 156; Mismatches 418; Indels 279; Gaps 36;



```
Qy        63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
Db       137 YMAAQENHEEVVKYLLKHGANQALSTEDGFTPLAVA-------------------LQQG 176

Qy       123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181
Db       177 HDRVVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTP 230

Qy       182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
Db       231 LHIAAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IID 287

Qy       242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301
Db       288 SRTKDLL-------------------------------------------------TP 296

Qy       302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
Db       297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy       362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410
Db       354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411

Qy       411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457
Db       412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy       458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517
Db       457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy       518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576
Db       516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy       577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633
Db       576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy       634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693
Db       611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670

Qy       694 LCALLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPS 739
Db       671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy       740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
Db       730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy       797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852
Db       779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

Qy       853 ----------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG 884
Db       836 ETMFSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGG 895

Qy       885 ------SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQ 932
Db       896 AIPKINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLH 950

Qy       933 KVCNKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DE 980
Db       951 GAGPEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDA 1010

Qy       981 RHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKS 1040
Db      1011 RGGAMRGCRHSGVRIIVPPRKASQ----------PIR-----------VTCRY-LRKD 1046

Qy      1041 FLQFSAMKMAHSPP 1054
```

```
Db      1047 -------KLAHPPP 1053
```

RESULT 26
Q17488
ID   Q17488 PRELIMINARY; PRT; 1815 AA.
AC   Q17488; O02517;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   C. ELEGANS ankyrin-related UNC-44 (GB:U21734).
GN   B0350.2 OR UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gattung S.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U50071; AAA93445.1; -.
DR   EMBL; U39847; AAB41826.1; -.
DR   HSSP; P42773; 1IHB.
DR   WormPep; B0350.2d; CE06705.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.

```
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1815 AA;  198956 MW;  3ECC6E310915C915 CRC64;

  Query Match           14.0%;  Score 810.5;  DB 5;  Length 1815;
  Best Local Similarity   27.3%;  Pred. No. 5.9e-43;
  Matches  273;  Conservative  139;  Mismatches  362;  Indels  225;  Gaps   27;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |::  | | |    |  || |   | : :| :|: :||: | :   :|||
Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
              : || :   :|   ||:|||   :   || | | :|                    : |
Db        137 YMAAQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQG 176

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181
              ::  :  ||       : | |:   | ||:||  :      ||||:  :      |    |
Db        177 HDRVVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTP 230

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
              || |  ||| | :||:: || ||       :||| |     |  : :|||| ||   :::
Db        231 LHIAAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IID 287

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301
              |  :                                                       |
Db        288 SRTKDLL-------------------------------------------------TP 296

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
              ||||| | :    |: :||: :|| |: |||  | |||:|::  | |    :: | | |:
Db        297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy        362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410
                :| |  | || ||:|||::  :|||    |||  :| ||| | :  :
Db        354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411

Qy        411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457
                   ::   : |::       :|        ||:     |||||
Db        412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517
              |  |||| ||  |:  ||  |:::|| | |: :    ||| |   |: ::  ||:: ||
Db        457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576
              |      ::||| || :|: |:  :|| | || |  : | ||| |   | |: :: |
Db        516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy        577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633
              |  ||                          | :|::  ||||:|| ||| :||  ||
Db        576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy        634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693
              ::||    |  | |  ||| ||    :::|: |::: |  ||   :  ||||| :||:|  :
Db        611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670

Qy        694 LCALLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPS 739
              :  ||: :|:|    |   | | : | : :|                :::   |   | : :
Db        671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy        740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
              |:  | || |:     ||      :|   :|    :|      |                |
Db        730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy        797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852
              ||  |:          || |    ||    :| |  : |   ||     |:  :|      :|
Db        779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

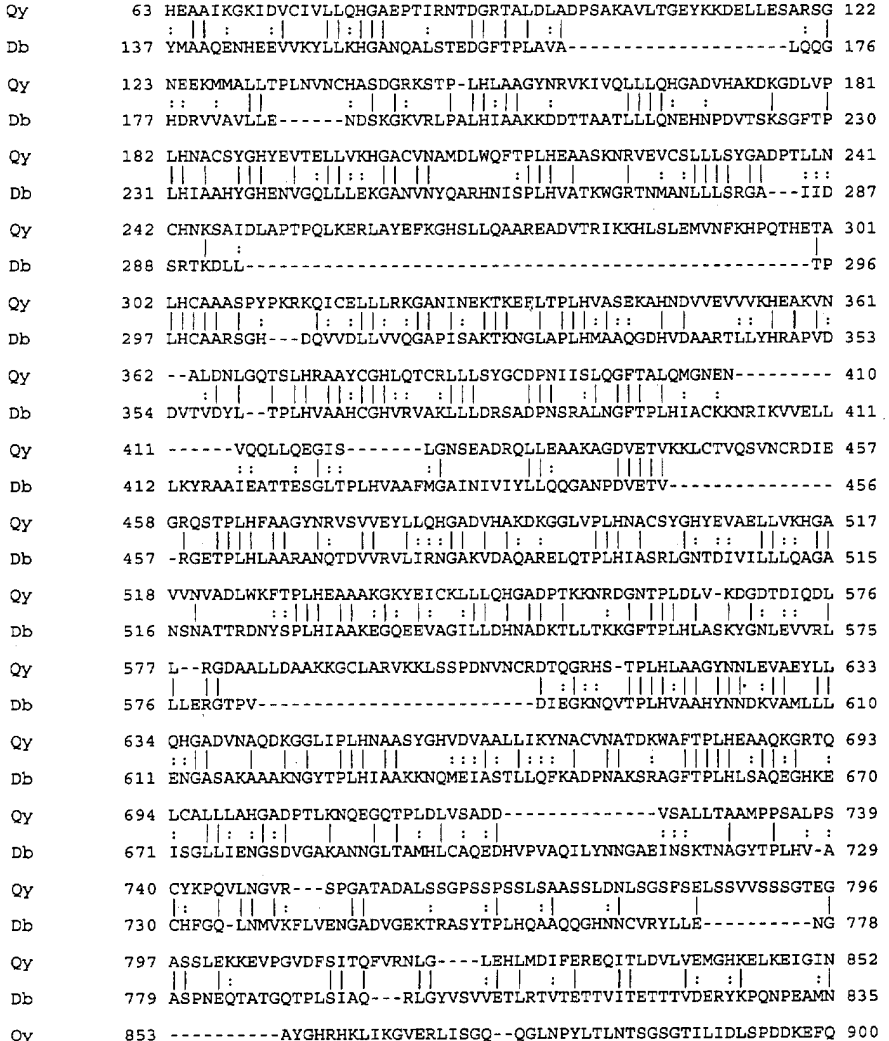
Qy        853 ----------AYGHRHKLIKGVERLISGQ--QGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
                         |: :   |:  |    |||          :|  :|
Db        836 ETMFSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQ-------DSTGVHMIHTG------- 881

Qy        901 SVEEEMQSTVREHRDGG-----HAGGIFNRYNILKIQKV 934
                |::    :|  :||      ::||:        ||  |
Db        882 ---EQLLQRSQELENGGAIPKINSGGMSPEKEFAKIAPV 917


RESULT 27
Q17343
ID   Q17343      PRELIMINARY;      PRT;  6994 AA.
AC   Q17343;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   UNC-44 ankyrins.
GN   UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
RN   [2]
RP   REVISIONS, AND SEQUENCE OF 6126-6994 FROM N.A.
RC   STRAIN=N2;
RA   Otsuka A.J.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U39847; AAB41827.1; -.
DR   EMBL; U21733; AAB38384.1; -.
DR   HSSP; P42773; 1IHB.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   6994 AA;  775364 MW;  90CB449925D9923D CRC64;

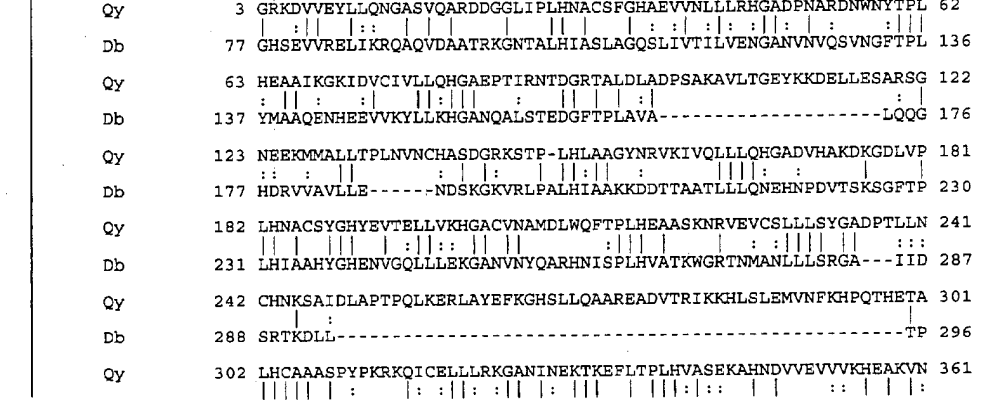
  Query Match           14.0%;  Score 810.5;  DB 5;  Length 6994;
  Best Local Similarity   27.3%;  Pred. No. 4.8e-42;
  Matches  273;  Conservative  139;  Mismatches  362;  Indels  225;  Gaps   27;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |::  | | |    |  || |   | : :| :|: :||: | :   :|||
Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136
```



RESULT 28
Q17490
ID Q17490 PRELIMINARY; PRT; 6994 AA.
AC Q17490;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Uncoordinated protein 44, isoform f.
GN B0350.2 OR UNC-44.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gattung S.;
RT "The sequence of C. elegans cosmid B0350.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U50071; AAA93447.2; -.
DR PIR; T15348; T15348.
DR WormPep; B0350.2f; CE31847.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR001360; Glyco_hydro_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 24.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 23.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 22.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
SQ SEQUENCE 6994 AA; 775382 MW; 0FFBBEEAD9408975 CRC64;

Query Match 14.0%; Score 810.5; DB 5; Length 6994;
Best Local Similarity 27.3%; Pred. No. 4.8e-42;
Matches 273; Conservative 139; Mismatches 362; Indels 225; Gaps 27;



```
Db        297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy        362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410

Db        354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411

Qy        411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457

Db        412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517

Db        457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576

Db        516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy        577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633

Db        576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy        634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693

Db        611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670

Qy        694 LCALLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPS 739

Db        671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy        740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796

Db        730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy        797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852

Db        779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

Qy        853 ----------AYGHRHKLIKGVERLISGQ--QGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Db        836 ETMFSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQ-------DSTGVHMIHTG------- 881

Qy        901 SVEEEMQSTVREHRDGG-----HAGGIFNRYNILKIQKV 934

Db        882 ---EQLLQRSQELENGGAIPKINSGGMSPEKEFAKIAPV 917
```

```
RESULT 29
Q17344
ID   Q17344      PRELIMINARY;      PRT;  1786 AA.
AC   Q17344;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   UNC-44 (Fragment).
GN   UNC-44.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2;
RX   MEDLINE=95263663; PubMed=7744957;
RA   Otsuka A.J., Franco R., Yang B., Shim K., Tang L.Z., Zhang Y.,
RA   Boontrakulpoontawee P., Jeyaprakash A., Hedgecock E., Wheaton V.I.,
RA   Sobery A.;
RT   "An ankyrin-related gene (unc-44) is necessary for proper axonal
RT   guidance in Caenorhabditis elegans.";
RL   J. Cell Biol. 129:1081-1092(1995).
DR   EMBL; U21734; AAA85854.1; -.
DR   HSSP; P42773; 1IHB.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000488; Death.
DR   InterPro; IPR001360; Glyco_hydro_1.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 24.
DR   Pfam; PF00531; death; 1.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 21.
DR   SMART; SM00005; DEATH; 1.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50017; DEATH_DOMAIN; 1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 2.
KW   ANK repeat; Repeat.
FT   NON_TER    1786   1786
SQ   SEQUENCE   1786 AA;  195550 MW;  03E220FB521747F1 CRC64;
```

Query Match 14.0%; Score 807.5; DB 5; Length 1786;
Best Local Similarity 26.0%; Pred. No. 8.9e-43;
Matches 302; Conservative 154; Mismatches 419; Indels 285; Gaps 35;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122

Db        137 YMAAQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQG 176

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181

Db        177 HDRVVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTP 230

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241

Db        231 LHIAAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IID 287

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301

Db        288 SRTKDLL-----------------------------------------------------TP 296

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361

Db        297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy        362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410

Db        354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411

Qy        411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457

Db        412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517

Db        457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576

Db        516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy        577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633

Db        576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy        634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693

Db        611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670
```

```
Qy        694 LCALLLAHGADPTLKNQEGQTPLDLVSADD--------------VSALLTAAMPPSALPS 739
Db        671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy        740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
Db        730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy        797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852
Db        779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

Qy        853 AY----------------GHRHKL------------IKGVERLISGQQGLNPYLTLNTSG 884
Db        836 ETMFSESEDEGQAAEHEGGGAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGG 895

Qy        885 ------SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQ 932
Db        896 AIPKINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLH 950

Qy        933 KVCNKKLWERYTHRRKE------------VSEENHNHANERMLFHGSPFVNAIIHKGF-- 978
Db        951 GAGPEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSKLQHRTFLI 1010

Qy        979 ----DERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCR 1034
Db       1011 SFLVDARGGAMRGCRHSGVRIIVPPRKASQ----------PIR------------VTCR 1047

Qy       1035 VTLGKSFLQFSAMKMAHSPP 1054
Db       1048 Y-LRKD-------KLAHPPP 1059
```

RESULT 30
Q8SWY2
ID Q8SWY2 PRELIMINARY; PRT; 1009 AA.
AC Q8SWY2; Q9VSA2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE GH01626p (CG7462 protein).
GN ANK2 OR CG7462.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY094999; AAM11327.1; -.
DR EMBL; AE003558; AAF50525.2; -.
DR FlyBase; FBgn0017645; Ank2.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 19.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 19.
DR PROSITE; PS50088; ANK_REPEAT; 17.
DR PROSITE; PS50297; ANK_REP_REGION; 1.

```
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1009 AA;  109919 MW;  3AD24CEF12A77303 CRC64;

  Query Match           13.7%;  Score 793;  DB 5;  Length 1009;
  Best Local Similarity  25.5%;  Pred. No. 3.2e-42;
  Matches  279;  Conservative  146;  Mismatches  397;  Indels  270;  Gaps   27;

Qy          4 RKDVVE---YLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              :|| |:    || |  :      |  ||| |  :|:  : |||:: ||| |     | :
Db         31 KKDDVKAATLLLDNDHNPDVTSKSGFTPLHIASHYGNQNIANLLIQKGADVNYSAKHNIS 90

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||| ||  || ::  :||: |     :  || | |  |                    ||
Db         91 PLHVAAKWGKTNMVSLLLEKGGNIEAKTRDGLTPLHCA-------------------AR 130

Qy        121 SGNEEKMMALL---TPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 177
              ||:|: :  ||    |::    |        |||:||     |   ::|| | | |
Db        131 SGHEQVVDMLLERGAPIS----AKTKNGLAPLHMAAQGEHVDAARILLYHRAPVDEVTVD 186

Qy        178 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 237
              |  || |   ||  | :||:    |  ||  |  ||||| |  ||::|  ||| :||
Db        187 YLTALHVAAHCGHVRVAKLLLDRNADANARALNGFTPLHIACKKNRLKVVELLLRHGASI 246

Qy        238 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 297
              :      : : :|       : |      |||      ||  ::
Db        247 SATTESGLTPLHVAAFMGCMNIVIY------LLQHDASPDVPTVR--------------- 285

Qy        298 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 357
              || || || :    :  |  :||| || :: : :|  ||||:||   : |:| ::::|
Db        286 GETPLHLAARA---NQTDIIRILLRNGAQVDARAREQQTPLHIASRLGNVDIVMLLLQHG 342

Qy        358 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQE 417
              |:|:|      |:|| ||  |  :   :|:  |   :  : :||
Db        343 AQVDATTKDMYTALHIAAKEGQDEVAAVLIENGAALDAATKKGF---------------- 386

Qy        418 GISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVE 477
                                                    |||| | |  : | :
Db        387 ----------------------------------------TPLHLTAKYGHIKVAQ 402

Qy        478 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 537
              |||  ||| |: | |: ||| || | : :|| ||:: ||  :         |||| || |
Db        403 LLLQKEADVDAQGKNGVTPLHVACHYNNQQVALLLLEKGASPHATAKNGHTPLHIAARKN 462

Qy        538 KYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL-RGDAALLDAAKKGCLARV 595
              : :|   ||::||    ::: | ||| |  ::|  :| :||    ||:   || |
Db        463 QMDIATTLLEYGALANAESKAGFTPLHLSSQEGHAEISNLLIEHKAAVNHPAKNGL---- 518

Qy        596 KKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAAS 655
                              ||:|| |  :|: ||| | ::||:::   | |  ||| |:
Db        519 ------------------TPMHLCAQEDNVNVAEILEKNGANIDMATKAGYTPLHVASH 559

Qy        656 YGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 715
              :|  ::   |::  | |:|     :|||| ||:|   :  ||| | |:   :   ||||
Db        560 FGQANMVRFLLQNGANVDAATSIGYTPLHQTAQQGHCHIVNLLLEHKANANAQTVNGQTP 619

Qy        716 LDLVS-------ADDVSALL----TAAMPPSA-------LPSCYKPQVLNGVRSPGATAD 757
              | :          | :  :     |||  |  |       |       ::      |  :
Db        620 LHIARKLGYISVLDSLKTITKEDETAAAPSQAEEKYRVVAPEAMHESFMSDSEEEGGEDN 679

Qy        758 ALSSGP---------------SSPSSLSAASSLD-NLSGSFSELSS---------VVSSS 792
              ||  |               | |  ::    :| |    :| :|          |:|
Db        680 MLSDQPYRYLTVDEMKSLGDDSLPIDVTRDERMDSNRMTQSAEYASGVPPTIGEEVISPH 739

Qy        793 GTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIG-- 850
              |:   |  | | ||      :: |                      || |   :|
Db        740 KTQVYGSSPKATVDGV------YIAN--------------------GSGHDEPPHVGRK 772

Qy        851 -----------INAYGH-----RHKLIKGVE-------------RLISGQQGLNPYLTLN 881
                        ::| |      ||  :: :               | :  |: ::|   :
Db        773 LSWKSFLVSFLVDARGGAMRGCRHSGVRMIIPSRSTCQPTRVTCRYVKPQRTMHPPQLME 832

Qy        882 TSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWE 941
                  : :::| |   :|      :   | |      |   :  :   : | :  | : |
Db        833 GEALASRVLELGPCSTKF------IGPVVME---VPHFASLRGKEREIIILRSDNGETWR 883

Qy        942 RYTHRRKE-----------------VSEENHNHANERMLFHGSPFVNAIIHKGFDERHA 983
                :|    |                   : |:  ||   | : :  |   |:: :   | ||
Db        884 EHTIDNSEEIIHDVLQQCFEPEEIAQLEEQAGNHVC-RFVTYDFPQYFAVVSRIRQEVHA 942

Qy        984 Y--IGGMFGAGI 993
                   |||  : :
Db        943 IGPEGGMVSSTV 954
```

```
RESULT 31
Q8N8A2
ID   Q8N8A2      PRELIMINARY;      PRT;   919 AA.
AC   Q8N8A2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein FLJ39767.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RA   Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA   Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA   Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA   Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA   Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA   Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA   Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK097086; BAC04946.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 26.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 26.
DR   PROSITE; PS50088; ANK_REPEAT; 22.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   919 AA;  99819 MW;  FF98E15F0E9C4E89 CRC64;

  Query Match           11.8%;  Score 678;  DB 4;  Length 919;
  Best Local Similarity  28.5%;  Pred. No. 7.1e-35;
  Matches  235;  Conservative  109;  Mismatches  304;  Indels  178;  Gaps   24;

Qy         11 LLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGK 70
              |:     |   |     ||| |   | ||:: ||:  ||  ||:||   |||| |
Db          2 LIHKTEDVNTLDSEKRTPLHVAAFLGDAEIIELLILSGARVNAKDNMWLTPLHRAVASRS 61

Qy         71 IDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMAL 130
                :  ||::| |:   |: : :| | :|  : |||                     |  :
Db         62 EEAVQVLIKHSADVNARDKNWQTPLHVA-AANKAV-------------------KCAEV 100

Qy        131 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGH 190
              : ||  : : ||    | || ||     |::| |||  ||:::| || |   || |   ||
Db        101 IIPLLSSVNVSDRGGRTALHHAALNGHVEMVNLLLAKGANINAFDKKDRRALHWAAYMGH 160

Qy        191 YEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL 250
              :|  ||: ||| |   |   :|||| |||  :: |   ||: | :   :| :  :|: :
Db        161 LDVVALLINHGAEVTCKDKKGYTPLHAAASNGQINVVKHLLNLGVEIDEINVYGNTALHI 220

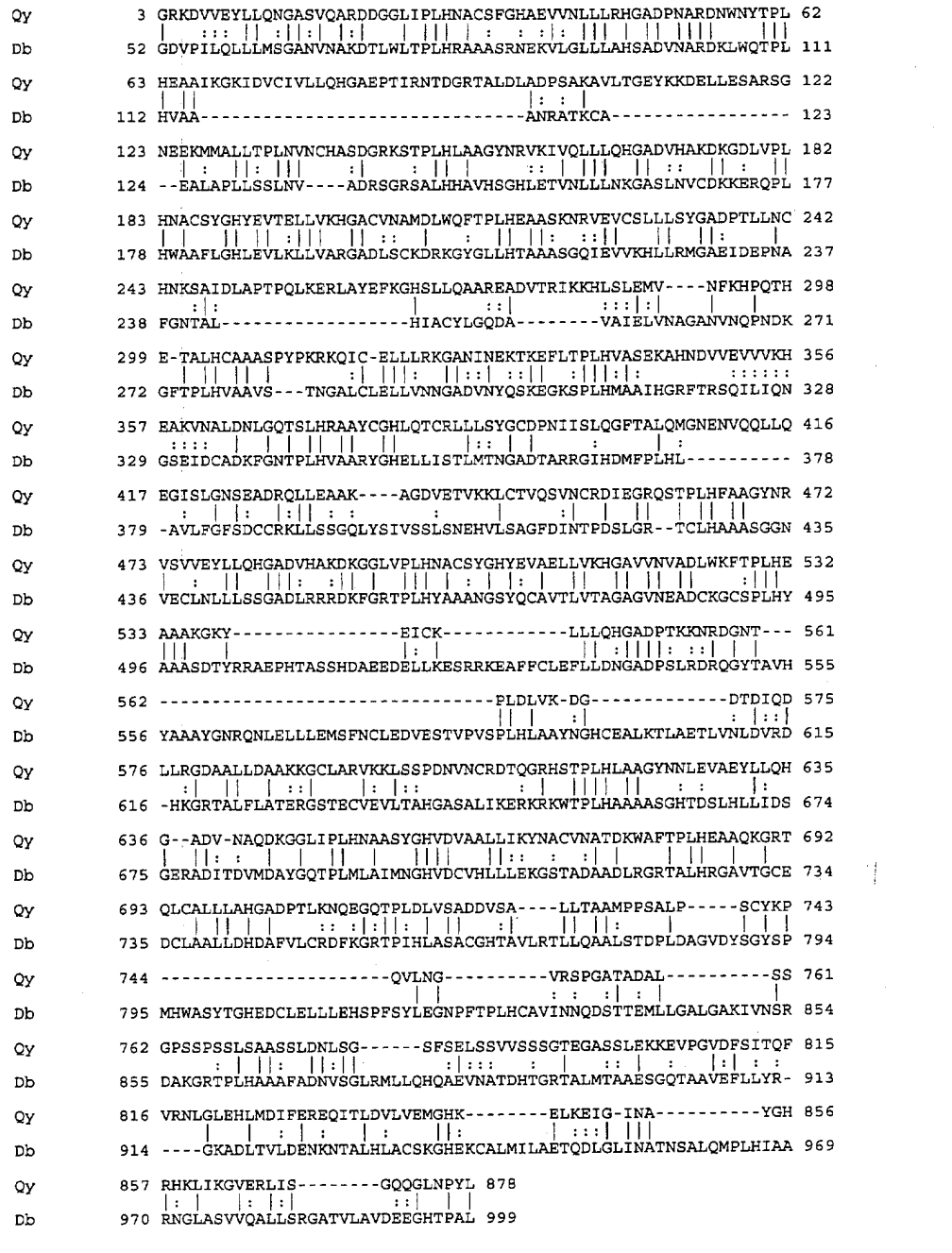
Qy        251 APTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHE-TALHCAAASP 309
              |          ::| | :  :     |   |  :  | || ||||
Db        221 A---------CYN---------GQDAVVNELIDY----GANVNQPNNNGFTPLHFAAAST 258

Qy        310 YPKRKQIC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQ 368
              :     :| |||:  ||::| ::|:  :|||: :         : :::: ::: :|  |
Db        259 H---GALCLELLVNNGADVNIQSKDGKSPLHMTAVHGRFTRSQTLIQNGGEIDCVDKDGN 315
```



RESULT 32
Q8BTI7
ID Q8BTI7 PRELIMINARY; PRT; 1076 AA.
AC Q8BTI7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical ankyrin-repeat containing protein.
GN G431002C21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK090124; BAC41105.1; -.
DR PIR; PT0649; PT0712.
DR PIR; PT0706; PT0706.
DR MGD; MGI:2444029; G431002C21Rik.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 28.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 28.
DR PROSITE; PS50088; ANK_REPEAT; 21.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW Hypothetical protein.
SQ SEQUENCE 1076 AA; 115031 MW; 058FE2C95E1A5682 CRC64;

Query Match 11.6%; Score 669; DB 11; Length 1076;
Best Local Similarity 25.6%; Pred. No. 3.5e-34;
Matches 269; Conservative 135; Mismatches 370; Indels 276; Gaps 31;



```
RESULT 33
Q8AX92
ID   Q8AX92      PRELIMINARY;      PRT;  1179 AA.
AC   Q8AX92;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kosan C., Kunz J.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF527754; AAO15006.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 23.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 22.
DR   PROSITE; PS50088; ANK_REPEAT; 16.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1179 AA;  127675 MW;  ABFF2F2B2CC14EF2 CRC64;

  Query Match           11.1%;  Score 640;  DB 13;  Length 1179;
  Best Local Similarity  27.6%;  Pred. No. 3e-32;
  Matches  252;  Conservative   99;  Mismatches  337;  Indels  226;  Gaps   29;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTP 61
              |  :::| |: :|| | |:|:  | ||| | :    : | :||:| || |||| ||  ||
Db        358 GDAEIIELLILSGARVNAKDNKWLTPLHRAVASCSEDAVAVLLKHSADVNARDKNWQ-TP 416

Qy         62 LHEAAIKGKIDVCIVL-----------------LQHGA-----EPTIRNTDGRTALDLAD 99
              || ||    :     |                 | | |     |   |  | | | ||
Db        417 LHVAASNKAVRCAEALVPLLSNVNVSDRAGRTALHHAAFSGHVEVNEVNAYGNTPLHLAC 476

Qy        100 PSAKAVLTGE---------------YKKDELLESARSGNEEKMMALLTPLNVNCHASDGR 144
               : : |: ||               :       |:| |    : : |    ::|  : | |
Db        477 YNGQDVVVGELIQAGAKVNQENERGFSPLHFASSSRQGALCQELLLTHGAHINMQSKDNR 536

Qy        145 KSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACV 204
               ||||:|| : |     | |:|:|||:  :||      || |    ||  :    |:|||
Db        537 --TPLHMAATHGRFSCSQALIQNGADIDCEDKNRNTALHIAARQGHELIITALIKHGANS 594

Qy        205 NAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEF 264
                  :   ||| ||     : |  ||| | |            ||   ||
Db        595 ARRGVHGMFPLHLAALSGFSDCCRKLLSSGFD-----------ID---TPD--------- 631

Qy        265 KGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKG 324
                                    :  | || |||      :    |||: |
Db        632 -------------------------------EIGRTCLHAAAAG---GNLECLNLLLKIG 657

Qy        325 ANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYC-----GH 379
              |: | |     |||| ||    :   |  :|   | || ||  |    || ||       ||
Db        658 ADFNRKDNFGRTPLHYASANCNYQCVPALVGSGASVNVLDQRGCNPLHYAAAADTEGKGH 717

Qy        380 LQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV 439
                 ||| :|  |          :    : :  :      |: |  | |: ::  ||
Db        718 QDCVALLLHHGASP----------MTRDYAHKKTAIHAAAMNGHQECLRLLMSHSQHLDV 767

Qy        440 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 499
               :              :|| |:  |||  |          | ||  || |  :|: |    ||
Db        768 D------------AQDINGQ--TPLMLAVLNGHTECVYSLLSQGASVENQDRWGRTALHR 813

Qy        500 ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTK----K 555
                 |  |  | |:: || | | |:   :||| |:| |:      |||  |  |
Db        814 GVVTGQEECVEALLQRGANVCVKDIQGRSPLHLASACGRVGALGALLQ--ASSTSHAHLT 871

Qy        556 NRDGNTPLD-LVKDG-----------DTDIQDLLRGD-------AALLDAAKKGCLARVK 596
```

```
Db        872 DNQGYTPLHWACYNGTGYDSCVEVLLDQEVFKQIKGNSFSPLHCAVIND--NEGVAEMLI 929

Qy        597 KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGA------------------ 637
                 :    : :|  |:||  |||| ||  :::|     || |||
Db        930 ESMGTNIINTSDSKGR--TPLHAAAFSDHVECVSLLLSHGAQANVADTRLCRTPLMMAAL 987

Qy        638 ---------------DVNAQDKGGLIPLHNAASYGHVDVAALLIKY---NACVNATDKW 678
                            |:  ||       || | | ||    | |:::       :| |:
Db        988 NGQTNTVEVLVNSAKVDLTLQDAHRNTALHLACSKGHETCALLILEKIGDRNLINCTNAA 1047

Qy        679 AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL-------DLVSADDVSALLTAA 731
               .|||| ||::| | :   ||    |      :: | ||         |:  || :: :| :
Db       1048 LQTPLHVAARRGLTVVVQELLGKEASVLAVDENGYTPALSCTPNRDV--ADCLALILNSM 1105

Qy        732 MPPS------ALPSCYKPQVL------NGVRSPGATADALSSGPSSPS---------SLS 770
               || |      |||:    | :       |  | |   |||   |  |:          ||
Db       1106 MPTSPMVTIAALPAISITQTVINHHPTNNHISKGVGFDALP--PLRPNHASYCRAERPLS 1163

Qy        771 AASSLDNLSGSFSE 784
               : |: : |: | ||
Db       1164 SVSADEELNDSDSE 1177


RESULT 34
Q7T1G6
ID   Q7T1G6      PRELIMINARY;      PRT;  1614 AA.
AC   Q7T1G6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ion channel NompC.
GN   NOMPC.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22728086; PubMed=12805553;
RA   Sidi S., Friedrich R.W., Nicolson T.;
RT   "NompC TRP channel required for vertebrate sensory hair cell
RT   mechanotransduction.";
RL   Science 301:96-99(2003).
DR   EMBL; AY313897; AAP86445.1; -.
SQ   SEQUENCE   1614 AA;  175798 MW;  773343F37E3CB2DD CRC64;

  Query Match           10.9%;  Score 631;  DB 13;  Length 1614;
  Best Local Similarity  24.3%;  Pred. No. 1.8e-31;
  Matches  259;  Conservative  157;  Mismatches  381;  Indels  270;  Gaps   31;

Qy          4 RKDV--VEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 61
              |:||   : |:: ||:  :::| |  ||| |    |   ::   |     |: |  |  : :|
Db        197 RRDVEMAKILVEFGANPDSQNDEGQTPLHIAAHEGDENMLKFLYLCKANANISDKMDRSP 256

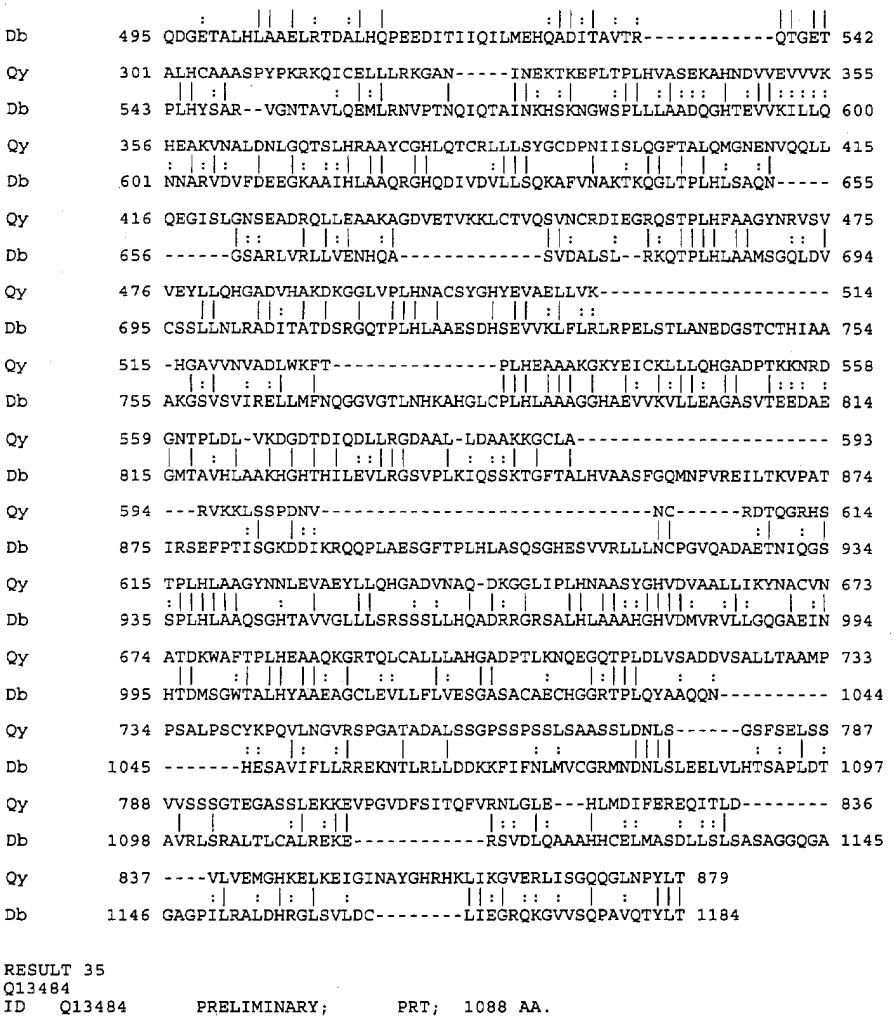
Qy         62 LHEAAIKGKIDVCIVLLQHGAEPTI-RNTDGRTALDLADPSAKAVLTGEYKKDE------ 114
              || || :|  :|  :| :       : |  || | | :|          : :
Db        257 LHIAAERGHTNVVEILTEKFRSCVLARTKDGNTLLHIASQCGHPTTALSFLRKGVPLHMP 316

Qy        115 -------LLESARSGNEEKMMALLTP-LNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQ 166
                    |  :|: |:   : |||    :|:  | ||:  | ||:|     | ::||:||
Db        317 NKSGAVCLHAAAKRGHTAVVKALLQKGAHVDAAARDGQ--TALHIAVENCRPQVVQMLLG 374

Qy        167 HGADVHAK-DKGDLVPLHNACSYGHYE-VTELLVKHGACVNAMDLWQFTPLHEAASKNRV 224
              || |  :  |     ||| :       |    |:|:| || ||| |       | || ||    :
Db        375 FGAHVQLRGGKAQETPLHISARVKEGERAAEMLLKSGAEVNAEQENGETALHVAARHGSL 434

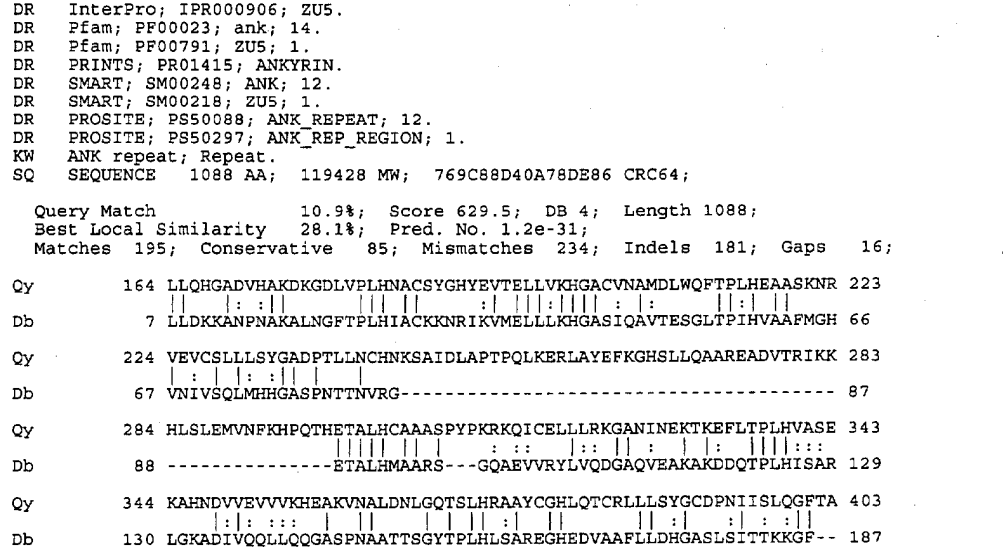
Qy        225 EVCSLLLSYGADP-------------TLLNCH-------------NKSAIDLA------- 251
               ::   |:  | ||               : :||             ||  |
Db        435 QMIRALIQEGGDPRWRSRVGESPLHVAVRHCHAHVVQEILTFLTNEKSRRDAELCVCEGN 494

Qy        252 PTPQLKERLAYEFKGHSLLQ----------AAREADVTRIKKHLSLEMVNFKHPQTHET 300
```



```
RESULT 35
Q13484
ID   Q13484      PRELIMINARY;      PRT;  1088 AA.
AC   Q13484;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ankyrin G119.
GN   ANK3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=96234670; PubMed=8666667;
RA   Devarajan P., Stabach P.R., Mann A.S., Ardito T., Kashgarian M.,
RA   Morrow J.S.;
RT   "Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney
RT   and muscle that binds beta I sigma spectrin and associates with the
RT   Golgi apparatus.";
RL   J. Cell Biol. 133:819-830(1996).
DR   EMBL; U43965; AAB08437.1; -.
DR   HSSP; P55273; 1BI8.
DR   GO; GO:0005783; C:endoplasmic reticulum; TAS.
DR   GO; GO:0005794; C:Golgi apparatus; TAS.
DR   GO; GO:0007016; P:cytoskeletal anchoring; TAS.
DR   GO; GO:0006899; P:nonselective vesicle transport; TAS.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000906; ZU5.
DR   Pfam; PF00023; ank; 14.
DR   Pfam; PF00791; ZU5; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 12.
DR   SMART; SM00218; ZU5; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 12.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1088 AA;  119428 MW;  769C88D40A78DE86 CRC64;

  Query Match             10.9%;  Score 629.5;  DB 4;  Length 1088;
  Best Local Similarity   28.1%;  Pred. No. 1.2e-31;
  Matches  195;  Conservative   85;  Mismatches  234;  Indels  181;  Gaps   16;
```



```
RESULT 36
Q9UFA4
ID   Q9UFA4       PRELIMINARY;      PRT;   791 AA.
AC   Q9UFA4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP434D2328.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
RA   Wiemann S.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL133087; CAB61404.1; -.
DR   PIR; T42691; T42691.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 22.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 19.
DR   PROSITE; PS50088; ANK_REPEAT; 17.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
FT   NON_TER       1       1
SQ   SEQUENCE   791 AA;  85632 MW;  43626C2F4DF558EF CRC64;

  Query Match             10.9%;  Score 628;  DB 4;  Length 791;
  Best Local Similarity   27.7%;  Pred. No. 9.5e-32;
  Matches  232;  Conservative  119;  Mismatches  313;  Indels  174;  Gaps   28;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||::||  |  :   :  |   || ||  |   ||| |: :||: |  :|  :||| 
Db         14 GQINVVKHLLNLGVEIDEINVYGNTALHIACYNGQDAVVNELIDYGANVNQPNNNGFTPL 73

Qy         63 HEAAIKGKIDVCI-VLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELL----- 116
              | ||      :|: :|: :||:  |:: ||:: | :       : | : : : |
Db         74 HFAAASTHGALCLELLVNNGADVNIQSKDGKSPLHM------TAVHGRFTRSQTLIQNGG 127

Qy        117 --------------ESARSGNEEKMMALLT--------------PLN---VNCHA----- 140
                           :|| |:|  :  |:|               ||:   :| |:
Db        128 EIDCVDKDGNTPLHVAARYGHELLINTLITSGADTAKCGIHSMFPLHLAALNAHSDCCRK 187

Qy        141 --SDGRK-------------------STP-------LHLAAGYNRVKIVQLLLQHGADVH 172
                 | |:|                  ||        || ||     |: ::||   ||| |
Db        188 LLSSGQKYSIVSLFSNEHVLSAGFEIDTPDKFGRTCLHAAAAGGNVECIKLLQSSGADFH 247

Qy        173 AKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLS 232
              |||     ||| | :  |:   | ||  || ||  | |  | || ||: :
Db        248 KKDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDDWGRTALHYAAASD---------- 297

Qy        233 YGADPTLL-NCHNKSAIDLAPTPQLKER---LAYEFKGHSLLQAAREADVTRIKKHLSLE 288
                 : |:| | |: |  :|    :|||:   |  ||    |||
Db        298 MDRNKTILGNAHDNSE-ELERARELKEKEATLCLEF----LLQ---------------- 335

Qy        289 MVNFKHPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT--PLHVAS 342
              |  :|   :     ::| |||  :   :|  |||| :  :  |::   |  |||:|:
Db        336 --NDANPSIRDKEGYNSIHYAAAYGH---RQCLELLLERTNSGFEESDSGATKSPLHLAA 390

Qy        343 EKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFT 402
               |:  :||:::    ::  |  |:|:|  ||: || :    |::           || :
Db        391 YNGHHQALEVLLQSLVDLDIRDEKGRTALDLAAFKGHTECVEALIN----------QGAS 440

Qy        403 ALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQST 462
                 |  :  |   :  |::   | ||| |              ::|: :| :|:  |
Db        441 IFVKDNVTKRTPLHASVINGHTLCLRLLLEIAD----------NPEAVDVKDAKGQ--T 487

Qy        463 PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 522
              || |  |  :  |  ||:  |:|   |  |  ||     || |  ::|::    :
Db        488 PLMLAVAYGHIDAVSLLLEKEANVDTVDILGCTALHRGIMTGHEECVQMLLEQEVSILCK 547

Qy        523 DLWKFTPLHEAAAKGKYEICKLLLQHG---ADPTKKNRDGNTPLD--LVKDGDTDIQDLL 577
              |    |||| |||:|        |||        |  |:  | |||          :  |: ||
Db        548 DSRGRTPLHYAAARGHATWLSELLQMALSEEDCCFKDNQGYTPLHWACYNGNENCIEVLL 607

Qy        578 R--------GD-------AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAG 622
                       |:       | : |      | : :        |:||| :||  |||| ||
Db        608 EQKCFRKFIGNPFTPLHCAIINDHG--NCASLLLGAIDSSIVSCRDDKGR--TPLHAAAF 663

Qy        623 YNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLI-KYNACVNATDKWAFT 681
               :::|  : ||:| | ||| |  |    |  || |      :|:    | :   ||  |
Db        664 ADHVECLQLLLRHSAPVNAVDNSGKTALMMAAENGQAGAVDILVNSAQADLTVKDKDLNT 723

Qy        682 PLHEAAQKGRTQLCALLLAHGADPTL---KNQEGQTPLDLVSADDVSALLTAAMPPSA 736
              ||| |  ||  :   |:|    | :|   ||   |||| : : : :  ::  :  :  |
Db        724 PLHLACSKGHEKCALLILDKIQDESLINEKNNALQTPLHVAARNGLKVVVEELLAKGA 781


RESULT 37
Q8IZ72
ID   Q8IZ72       PRELIMINARY;      PRT;   638 AA.
AC   Q8IZ72;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Duodenum;
RA   Strausberg R.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC016985; AAH16985.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 16.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 16.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   638 AA;  69275 MW;  C1C4E40F7F21B13F CRC64;

  Query Match             10.5%;  Score 606;  DB 4;  Length 638;
  Best Local Similarity   28.2%;  Pred. No. 1.8e-30;
  Matches  203;  Conservative   92;  Mismatches  256;  Indels  170;  Gaps   18;

Qy         32 ACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG 91
                | || ::: :|||   | |:|      :|       |:   |:   : |    :: ||
Db         19 AVKFG-SQIFSLLL---AIPSA------SPAPGVTSAGRGPACVCARRSGRMAVLKLTD- 67

Qy         92 RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHL 151
                        | :|: :|:                ||  | :    :||    |  | ||||:
Db         68 ------QPPLVQAIFSGD------------PEEIRMLIHKTEDVN--TLDSEKRTPLHV 106

Qy        152 AAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDL-W 210
              ||      :|::||:  || |:|||    | ||| | :     |  ::|:|| | ||| |  |
Db        107 AAFLGDAEIIELLILSGARVNAKDNMWLTPLHRAVASRSEEAVQVLIKHSADVNARDKNW 166

Qy        211 QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL 270
              | |||| ||:   |:   :::                                      ||
Db        167 Q-TPLHVAAANKAVKCAEVII-----------------------------------PLL 189

Qy        271 QAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEK 330
               :  :|                  :   |||| || : :    ::  ||| ||||||
Db        190 SSVNVSD-----------------RGGRTALHHAALNGH---VEMVNLLLAKGANINAF 228

Qy        331 TKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYG 390
```

```
Db        229 DKKDRRALHWAAYMGHLDVVALLINHGAEVTCKDKKGYTPLHAAASNGQINVVKHLLNLG 288

Qy        391 CDPNIISLQGFTALQMGNENVQQ-LLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQ 449

Db        289 VEIDEINVYGNTALHIACYNGQDAVVNELIDYG-------------------------- 321

Qy        450 SVNCRDIEGRQSTPLHFAAGYNRVSV-VEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 508

Db        322 -ANVNQPNNNGFTPLHFAAASTHGALCLELLVNNGADVNIQSKDGKSPLHMTAVHGRFTR 380

Qy        509 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKD 568

Db        381 SQTLIQNGGEIDCVDKDGNTPLHVAARYGHELLINTLITSGADTAKCGIHSMFPLHL--- 437

Qy        569 GDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEV 628

Db        438 ------------AALNAHSDCCR---KLLSSGFEIDTPDKFGR--TCLHAAAAGGNVEC 479

Qy        629 AEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA- 687

Db        480 IKLLQSSGADFHKKDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDDWGRTALHYAAA 539

Qy        688 --------------------------QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        540 SDMDRNKTILGNAHDNSEELERARELKEKEATLCLEFLLQNDANPSIRDKEGYNSIHYAA 599

Qy        721 A 721
              |
Db        600 A 600
```

```
RESULT 38
Q9GRV5
ID   Q9GRV5      PRELIMINARY;      PRT;  1913 AA.
AC   Q9GRV5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Y71A12B.4 protein.
GN   Y71A12B.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sulston J.E.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; AL132902; CAC14420.1; -.
DR   HSSP; Q00420; 1AWC.
DR   WormPep; Y71A12B.4; CE26238.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005262; F:calcium channel activity; IEA.
DR   GO; GO:0006816; P:calcium ion transport; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002111; Cat_channel_TrpL.
DR   InterPro; IPR002153; Trans_receptor.
DR   Pfam; PF00023; ank; 29.
DR   PRINTS; PR01415; ANKYRIN.
DR   PRINTS; PR01097; TRNSRECEPTRP.
DR   SMART; SM00248; ANK; 24.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Ionic channel; Repeat; Transmembrane.
```

```
SQ   SEQUENCE   1913 AA;  211430 MW;  D421B7F253C42A27 CRC64;

  Query Match             10.4%;  Score 601;  DB 5;  Length 1913;
  Best Local Similarity   27.3%;  Pred. No. 2.1e-29;
  Matches  242;  Conservative  113;  Mismatches  338;  Indels  194;  Gaps   30;

Qy          7 VVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR--DNWNYTPLHE 64

Db        392 IVKRILEDTDQRMAEDGDGSLPIHLAFKFGNVNIVELLLSGPSDEQTRKADGNGDTLLHL 451

Qy         65 AAIKGKIDVCIVLLQHGAE-PTIRNTDGRTAL----DLADPS---------AKAVLTGEY 110

Db        452 AARSGNIEAVRTAIAAGCDNANVQNRVGRTPLHEVAEVGDQNMLKIMFKLRADANIHDKE 511

Qy        111 KKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGAD 170

Db        512 DKTPVHVAAERGDTSMVESLIDKFGGSIRARTRDGSTLLHIAACSGHTSTALAFLKRGVP 571

Qy        171 VHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLL 230

Db        572 LFMPNKKGALGLHSAAAAGFNDVVKMLIARGTNVDVRTRDNYTALHVAVQSGKASVVETL 631

Qy        231 LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS-------LLQAAREADVTRIKK 283

Db        632 LGSGAD-----IHVKGG-ELGQTAL---HIAASLNGAESRDCAMMLLKSGGQPDVAQM-- 680

Qy        284 HLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASE 343

Db        681 ------------DGETCLHIAARS---GNKDIMRLLLDENADSKISSKIGETPLQVAAK 724

Qy        344 KAHNDVVEVVVKH----------EAKVNALDNLGQTSLHRAAYCGHLQ---------TCR 384

Db        725 SCNFEAASMILKHLSEVLTQEQLKEHVNHRTNDGFTALHYAAEIEQRQLHFPGEDAKLVN 784

Qy        385 LLLSYGCDPNIISLQ-GFTALQM----GNENVQQLLQEGISLG------NSEADR---QL 430

Db        785 LLIDYGGMVEMPSLNANETAMMHAARSGNQAVLLAMVNKIGAGAVQIVQNKQSKNGWSPL 844

Qy        431 LEAAKAGDVETVKKLCTVQS-VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAK 489

Db        845 LEACARGHSGVANILLKHHARIDVFDEMGR--TALHLAAFNGHLSLVHLLLQHKAFVNSK 902

Qy        490 DKGGLVPLHNACSYGHYEVAELLVK-HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQH 548

Db        903 SKTGEAPLHLAAQHGHVKVVNVLVQDHGAALEAITLDNQTALHFAAKFGQLAVSQTLLAL 962

Qy        549 GADPTKKNRDGNTPLDLVKDGD-TDIQDLL-------RGDAALLD---------AAKKGC 591

Db        963 GANPNARDDKGQTPLHLAAENDFPDVVKLFLKMRNNNRSVLTAIDHNGFTCAHIAAMKGS 1022

Qy        592 LARVKKLSSPDN--VNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQD------ 643

Db       1023 LAVVRELMMIDKPMVIQAKTKTLEATTLHMAAAGGHANIVKILLENGA--NAEDENSHGM 1080

Qy        644 ----------------------------KGGLIPLHNAASYGHVDVAALLIK-------- 667

Db       1081 TALHLGAKNGFISILEAFDKILWKRCSRKTGLNALHIAAFYGNSDFVNEMLKHVQATVRS 1140

Qy        668 ----YNACVNA--TDKWAFTPLHEAAQKGRTQLCALLLAHG--ADPT------------- 706

Db       1141 EPPIYNHHVNKEFSTEYGFTPLHLAAQSGHDSLVRMLLNQGVQVDATSTTMNVIPLHLAA 1200

Qy        707 --------------------LKNQEGQTPLDLVSAD---DVSALLTA 730

Db       1201 QQGHIAVVGMLLSRSTQQQHAKDWRGRTPLHLAAQNGHYEMVSLLIA 1247
```

```
RESULT 39
Q9VMR4
ID   Q9VMR4      PRELIMINARY;      PRT;  1619 AA.
AC   Q9VMR4; Q9NGR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
```

```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG11020 protein (Mechanosensory transduction channel NOMPC).
GN   NOMPC OR CG11020 OR CG17468.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   Walker R.G., Willingham A.T., Zuker C.S.;
RT   "A Drosophila mechanosensory transduction channel.";
RL   Science 0:0-0(2000).
DR   EMBL; AE003609; AAF52248.3; -.
DR   EMBL; AF242296; AAF59842.1; -.
DR   HSSP; Q00420; 1AWC.
DR   FlyBase; FBgn0016920; nompC.
DR   GO; GO:0007605; P:hearing; IMP.
DR   GO; GO:0007638; P:mechanosensory behavior; IMP.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR002111; Cat_channel_TrpL.
DR   InterPro; IPR000345; CytC_heme_BS.
DR   InterPro; IPR000550; Hppk.
DR   Pfam; PF00023; ank; 29.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 28.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS00190; CYTOCHROME_C; 1.
DR   PROSITE; PS00794; HPPK; 1.
KW   ANK repeat; Ionic channel; Repeat; Transmembrane.
SQ   SEQUENCE   1619 AA;  176844 MW;  B4FE239ABB1D1138 CRC64;

  Query Match           10.1%;  Score 584.5;  DB 5;  Length 1619;
  Best Local Similarity  26.2%;  Pred. No. 1.9e-28;
  Matches  237;  Conservative  118;  Mismatches  318;  Indels  231;  Gaps   29;

Qy          3 GRKDVVEYLL--QNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              | : :    ||  |    ::|     |   || |       ::| :|: :| : : ::     |
Db        244 GNQSMCRELLAAQTAEQLKATTANGDTALHLAARRRDVDMVRILVDYGTNVDTQNGEGQT 303

Qy         61 PLHEAAIKGKIDVCIVLLQHG--AEPTIRNTDGRTALDLADPSAKA----VLTGEYKKDE 114
              ||| || :|  |  ::   :|  |  :| :  || :| |  :  |    :|  ::|
Db        304 PLHIAAAEG--DEALLKYFYGVRASASIADNQDRTPMHLAAENGHAHVIEILADKFKA-S 360

Qy        115 LLESARSGNEEKMMALLTPLNVNCHA----------------SDGRKSTPLHLAAGYNR 157
              : |  : |:       |:   ::| ||                || :|  :| || |
Db        361 IFERTKDGS-----TLMHIASLNGHAECATMLFKKGVYLHMPNKDGARS--IHTAAAYGH 413

Qy        158 VKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM-DLWQFTPLH 216
              |:  ||| |  |       :   || |       | | |:  || |:      : ||||
Db        414 TGIINTLLQKGEKVDVTTNDNYTALHIAVESAKPAVVETLLGFGADVHVRGGKLRETPLH 473

Qy        217 EAASKNRVEVCSL-LLSYGADPTLL--NCHNKSAI-----DLAPTPQLKE----RLAYEF 264
              ||      : |:| ||  || | |   :|      :     :||   || |      |
Db        474 IAARVKDGDRCALMLLKSGASPNLTTDDCLTPVHVAARHGNLATLMQLLEDEGDPLYKSN 533

Qy        265 KGHSLLQAAREADVTRIKKHLSLEMVNFKH-----------PQTHETALH--CAAASPY 310
              | : |  |  |  | :|| : | |  ||                :   |||| |
Db        534 TGETPLHMACRACHPDIVRHL-IETVKEKHGPDKATTYINSVNEDGATALHYTCQITKEE 592

Qy        311 PK----RKQICELLLRKGANINEKTKEPL---------------------------- 335
              |     |||  :|| ||::  :|| |
Db        593 VKIPESDKQIVRMLLENGADVTLQTKTALETAFHYCAVAGNNDVLMEMISHMNPTDIQKA 652

Qy        336 ---------TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
                      ||| :|  :  | ::|  :: : |:|:  |  |:::|| ||  |:|  |  |
Db        653 MNRQSSVGWTPLLIACHRGHMELVNNLLANHARVDVFDTEGRSALHLAAERGYLHVCDAL 712
```

```
Qy        387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
              |:     |  |  | ||| :                          ||  |    || |
Db        713 LTNKAFINSKSRVGRTALHL--------------------------AAMNGFTHLVKFLI 746

Qy        447 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 506
                :    :  |: |||| ||   :: | : ||: ||:: | |  |  |:| |    :
Db        747 KDHNAVIDILTLRKQTPLHLAAASGQMEVCQLLLELGANIDATDDLGQKPIHVAAQNNYS 806

Qy        507 EVAEL-LVKHGAVVN----------------------VADLWKF---------------TP 529
              |||:| | :| ::||                      : :| ||               ||
Db        807 EVAKLFLQQHPSLVNATSKDGNTCAHIAAMQGSVKVIEELMKFDRSGVISARNKLTDATP 866

Qy        530 LHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAK 588
              |  ||  |  :: | |:: ||  |::|: | | : |  ::|   : |:|:  :|   :|
Db        867 LQLAAEGGHADVVKALVRAGASCTEENKAGFTAVHLAAQNGHGQVLDVLKSTNSLRINSK 926

Qy        589 K------------GCLARVKKL--SSPDNVNCRDTQGRH----------STPLHLAA--- 621
              |            |    |::|  | |  |        |:           |||||||
Db        927 KLGLTPLHVAAYYGQADTVRELLTSVPATVKSETPTGQSLFGDLGTESGMTPLHLAAFSG 986

Qy        622 ----------------------GYNNLEVAEY---------LLQHGAD-VNAQDKGGLI 648
                                   ||| | :| :          ||   |: : :||: |
Db        987 NENVVRLLLNSAGVQVDAATIENGYNPLHLACFGGHMSVVGLLLSRSAELLQSQDRNGRT 1046

Qy        649 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 708
              || || :||: :  :|:   | :||||:  :|||| ||: |  ::  ||   || |  :
Db       1047 GLHIAAMHGHIQMVEILLGQGAEINATDRNGWTPLHCAAKAGHLEVVKLLCEAGASPKSE 1106

Qy        709 NQEG 712
                |
Db       1107 TNYG 1110
```

RESULT 40
Q61304
ID Q61304 PRELIMINARY; PRT; 1098 AA.
AC Q61304;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Erythroid ankyrin (Fragment).
GN ANK1 OR ANK-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Spleen;
RX MEDLINE=93252825; PubMed=8486643;
RA Birkenmeier C.S., White R.A., Peters L.L., Hall E.J., Lux S.E.,
RA Barker J.E.;
RT "Complex patterns of sequence variation and multiple 5' and 3' ends
RT are found among transcripts of the erythroid ankyrin gene.";
RL J. Biol. Chem. 268:9533-9540(1993).
DR EMBL; X69065; CAA48803.1; -.
DR PIR; S37773; S37773.
DR HSSP; Q00420; 1AWC.
DR MGD; MGI:88024; Ank1.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00023; ank; 12.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 11.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50088; ANK_REPEAT; 11.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW ANK repeat; Repeat.
FT NON_TER 1 1
SQ SEQUENCE 1098 AA; 120875 MW; A67D1632A97C8B01 CRC64;

Query Match 9.6%; Score 554.5; DB 11; Length 1098;
Best Local Similarity 32.8%; Pred. No. 8.7e-27;
Matches 144; Conservative 74; Mismatches 158; Indels 63; Gaps 10;

```
Qy        300 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 359
              | || |:   :      | : ||::||: |     :  ||||:|:   | :| : :::::||
Db          3 TPLHVAS---FMGHLPIVKNLLQRGASPNVSNVKVETPLHMAARAGHTEVAKYLLQNKAK 59

Qy        360 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 419
               ||      || ||  ||     :|||  |  ||: :  | | |
Db         60 ANAKAKDDQTPLHCAARIGHTGMVKLLLENGASPNLATTAGHTPLH---------- 105

Qy        420 SLGNSEADRQLLEAAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                          ||: | |:|    |   : |   |   :|   |||| || | :| : |
Db        106 ------------TAAREGHVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVRLAEL 151

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEAAAKG 537
              ||:| |  :|  | || ||| |  : : :: :||:  |   : :  |  :|||| || :
Db        152 LLEHDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIAAKQN 210

Qy        538 KYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVK 596
              : |: : |||:|      ::  | ||| |  ::| |::  ||
Db        211 QIEVARSLLQYGGSANAESVQGVTPLHLAAQEGHTEMVALL------------------ 251

Qy        597 KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 656
               ||   | |  :  |  |||||  :   :: ||: |::||  |:|  :  |  ||| |: |
Db        252 -LSKQANGNLGNKSG--LTPLHLVSQEGHVLVADVLIKHGVTVDATTRMGYTPLHVASHY 308

Qy        657 GHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 716
              |:: :   |::: | |||   |  ::|||:|||:|  | :  ||| :|| |   :  | |||
Db        309 GNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNGTTPL 368

Qy        717 DL------VSADDVSALLT 729
                    :|  ||  ::|
Db        369 AIAKRLGYISVTDVLKVVT 387
```

RESULT 41
Q8C036
ID Q8C036 PRELIMINARY; PRT; 531 AA.
AC Q8C036;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical ankyrin-repeat containing protein.
GN G431002C21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK032429; BAC27865.1; -.
DR PIR; PT0649; PT0712.
DR PIR; PT0706; PT0706.
DR MGD; MGI:2444029; G431002C21Rik.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 15.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 14.
DR PROSITE; PS50088; ANK_REPEAT; 11.

```
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   531 AA;  57007 MW;  8CF4FFA30CDF48AC CRC64;

  Query Match             9.6%;  Score 551.5;  DB 11;  Length 531;
  Best Local Similarity   31.8%;  Pred. No. 4.4e-27;
  Matches  183;  Conservative   78;  Mismatches  222;  Indels   93;  Gaps   15;

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |:::  | : |:: :||:    | :  |  : |||| ||    | |:|||  ||:|:||
Db         12 LVQAIFSRDVEEVRSLLSQ-KENINVLDQERRTPLHAAAYVGDVPILQLLLMSGANVNAK 70

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMD-LWQFTPLHEAASKNRVEVCSLLLSY 233
              |   | ||| | :  : :|  ||: | | ||| | ||| |||| ||: ||   |
Db         71 DTLWLTPLHRAAASRNEKVLGLLLAHSADVNARDKLWQ-TPLHVAAA-NRATKC------ 122

Qy        234 GADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFK 293
                        | |||                || :   ||
Db        123 -------------AEALAP---------------LLSSLNVAD---------------- 137

Qy        294 HPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVV 353
                ::  :||| |  | :    :  ||| |||::|   |:  ||| |:   | :|::::
Db        138 --RSGRSALHHAVHSGH---LETVNLLLNKGASLNVCDKKERQPLHWAAFLGHLEVLKLL 192

Qy        354 VKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQ 413
              |  | ::  |  |  || || | ::  : || | : :  : | ||| :
Db        193 VARGADLSCKDRKGYGLLHTAAASGQIEVVKHLLRMGAEIDEPNAFGNTALHI------- 245

Qy        414 LLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAA-GYNR 472
                    ||    |  :|  ||           |     :  ||||  ||   |
Db        246 ----ACYLGQ---DAVAIELVNAG------------ANVNQPNDKGFTPLHVAAVSTNG 285

Qy        473 VSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHE 532
                :| |: :|||| : | |  ||| |  :| :  :::|:::|: :: || :  ||||
Db        286 ALCLELLVNNGADVNYQSKEGKSPLHMAAIHGRFTRSQILIQNGSEIDCADKFGNTPLHV 345

Qy        533 AAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGC 591
              ||  |   :  |: :|||  ::      || | |  | :|   |      |
Db        346 AARYGHELLISTLMTNGADTARRGIHDMFPLHLAVLFGFSDCCRKLLSSGQLYSIVSS-- 403

Qy        592 LARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLH 651
              |:    ||:  ::|  |: ||  | || ||   |:|    ||  |||:  :|| |  |||
Db        404 LSNEHVLSAGLDINTPDSLGR--TCLHAAASGGNVECLNLLLSSGADLRRRDKFGRTPLH 461

Qy        652 NAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA 687
               ||: |   |  |:   | ||  |    :||| ||
Db        462 YAAANGSYQCAVTLVTAGAGVNEADCKGCSPLHYAA 497


RESULT 42
Q02989
ID   Q02989      PRELIMINARY;      PRT;  1411 AA.
AC   Q02989;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha-latroinsectotoxin (Fragment).
OS   Latrodectus tredecimguttatus (black widow spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Araneoidea; Theridiidae; Latrodectus.
OX   NCBI_TaxID=6925;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=tredecimguttatus;
RX   MEDLINE=93238678; PubMed=8477689;
RA   Kiyatkin N.N.I., Dulubova I.I.E., Grishin E.E.V.;
RT   "Cloning and Structural Analysis of (alpha)-Latroinsectotoxin cDNA:
RT   Abundance of Ankyrin-Like Repeats.";
RL   Eur. J. Biochem. 213:121-127(1993).
DR   EMBL; Z14086; CAA78464.1; -.
DR   PIR; S30355; S30355.
DR   HSSP; P42773; 1IHB.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 18.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 15.
DR   PROSITE; PS50088; ANK_REPEAT; 10.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
FT   NON_TER       1      1
FT   CHAIN        36   1411       ALPHA-LATROINSECTOTOXIN.
SQ   SEQUENCE   1411 AA;  157805 MW;  B9001D1A5B0B7EF2 CRC64;

  Query Match             9.5%;  Score 545.5;  DB 5;  Length 1411;
  Best Local Similarity   23.7%;  Pred. No. 4.9e-26;
  Matches  239;  Conservative  136;  Mismatches  378;  Indels  257;  Gaps   33;

Qy         11 LLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLR------HGADPNARDNWNYTPLHE 64
              |:||||:|    : |   :| | | |: :|  |||         || |      |||||
Db        486 LIQNGANVSETFELGRGAIHAAASAGNYDVGELLLNKDINLLEKADKN-----GYTPLHI 540

Qy         65 AAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNE 124
              ||   | |  : |: : |:  :     ||  ||  |                 | :||
Db        541 AADSNKNDFVMFLIGNNADVNV-----RTKSDLFTP-------------LHLAARRDLT 581

Qy        125 EKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 184
              :   |:    :: :| |    |||||:          |:    | :: | |  | |||
Db        582 DVTQTLIDITEIDLNAQDKSGFTPLHLSISSTSETAAILIRNTNAVINIKSKVGLTPLHL 641

Qy        185 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 244
              |    :  |::||   || :|  |    ||||  ||    :|:   ||           |
Db        642 ATLQNNLSVSKLLAGKGAYLNDGDANGMTPLHYAAMTGNLEMVDFLL-----------N 689

Qy        245 KSAIDLAPTPQLKE----RLAYEFKGHSLLQ---------------------------- 271
              :  |::    : |:     ||  || : : :
Db        690 QQYININAATKEKKWTPLHLAILFKKNDVAERLLSDENLNIRLETNGGINPLHLASATGN 749

Qy        272 -------AAREADVTRI-KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK 323
                    |: |||||:  |  |                  |||         | ::|   |: |
Db        750 KQLVIELLAKNADVTRLTSKGFS-------------ALHLGIIG---KNEEIPFFLVEK 792

Qy        324 GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTC 383
              |||:|:||   :|||| |:     ::  :::   | : | |   |  :| |   |||:
Db        793 GANVNDKTNSGVTPLHFAAGLGKANIFRLLLSRGADIKAEDINSQMPIHEAVSNGHLEIV 852

Qy        384 RLLLSYGCDPNIISLQGFTALQMGNENVQQLLQE-GISLGNSEADRQLLEAAKAGDVETV 442
              |:|:    ||:::          ||: :  |    :  : : :  :  |
Db        853 RILIEK--DPSLM-------------NVKNIRNEYPFFYLAVEKRYKDIFDYFVSKD---- 893

Qy        443 KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKD------------ 490
                  | :::   :| ||  :   : ||::|:|:||:   |:
Db        894 --------ANVNEVDHNGNTLLHLFSSTGELEVVQFLMQNGANFRLKNNERKTFFDLAIE 945

Qy        491 ----------------------KGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 528
                                   :|  :  |  |    |:  |::      :| ::
Db        946 NGRLNIVAFAVEKNKVNLQAAHRGKTILYHAICDSAKYDKIEIVKYFIEKLNESEC---N 1002

Qy        529 PLHEAAAKGKYEICKLLLQ-HGADPTKKNRDGN-TPLDLVKDG-------DTDIQDLLRG 579
              |||||||    :: |  :|  | :| : |  :|  :   |       | |  | |
Db       1003 PLHEAAAYAHLDLVKYFVQERGINPAEFNEENQASPFCITIHGAPCGYSLDCDTPDRLEV 1062

Qy        580 DAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGAD 638
               | |                || |  | | |    :||: :|    | : :  ||:  |||
Db       1063 VEYLSD-------------KIPDINGKC-DVQ--ENTPITVAIFANKVSILNYLVGIGAD 1106

Qy        639 VNAQDKGGLIPLHNAASYGHVDVAALLIK-YNACVNATDKWAFTPLHEAAQKGRTQLCAL 697
               | |   |  ||: ||  |  ::   ||: :   :|  :|  || || ||:     :
Db       1107 PNQQVDGD-PPLYIAARQGRFEIVRCLIEVHKVDINTRNKERFTALHAAARNDFMDVVKY 1165

Qy        698 LLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATAD 757
              |:  |||   |  :   |:| :: :   | | ::      | | :  |  | : | | |
Db       1166 LVRQGADVNAKGIDDLRPID-IAGEKAKAYLQSS---RFLRSGHSFQ-SNEIDSFGNTIH 1220

Qy        758 ALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVR 817
               :|          :| : | |:   |   :   |: |||    |   | :| :
Db       1221 GIS---------MSARTNDKLTQQISSKGTRSDSNSTEGKMHSENVHVRSIDVN------ 1265
```

```
Qy        818 NLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHK------------LIKGV 864
Db       1266 --GALLLLDFMIRVFAS------------KKTNFAPYGSRIKTRSAAEAQAEALIMTERF 1311

Qy        865 ERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHR 914
Db       1312 ENLLSGLIG-----------------DPIPDSIDFSNVHSKIYKAIMSGR 1344


RESULT 43
O75179
ID   O75179      PRELIMINARY;      PRT;  2486 AA.
AC   O75179;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein KIAA0697 (Fragment).
GN   KIAA0697.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98403880; PubMed=9734811;
RA   Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA   Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. X.
RT   The complete sequences of 100 new cDNA clones from brain which can
RT   code for large proteins in vitro.";
RL   DNA Res. 5:169-176(1998).
CC   -!- SIMILARITY: CONTAINS 1 KH DOMAIN.
DR   EMBL; AB014597; BAA31672.2; -.
DR   PIR; T00353; T00353.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003684; F:damaged DNA binding; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0006298; P:mismatch repair; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000432; MutS_C.
DR   Pfam; PF00023; ank; 25.
DR   Pfam; PF00013; KH; 1.
DR   Pfam; PF00488; MutS_V; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 25.
DR   SMART; SM00322; KH; 1.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
DR   PROSITE; PS50084; KH_TYPE_1; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
FT   NON_TER      1      1
SQ   SEQUENCE   2486 AA;  263244 MW;  217C5B039240EEA9 CRC64;

  Query Match             9.3%;  Score 538;  DB 4;  Length 2486;
  Best Local Similarity   23.9%;  Pred. No. 3.6e-25;
  Matches  256;  Conservative  117;  Mismatches  363;  Indels  336;  Gaps   30;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDN------ 56
Db        229 GYVDVVKVLLESGASIEDHNENGHTPLMEAGSAGHVEVARLLLENGAGINTHSNEFKESA 288

Qy         57 --------------------------WNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 88
Db        289 LTLACYKGHLEMVRFLLEAGADQEHKTDEMHTALMEACMDGHVEVARLLLDSGAQVNMPA 348

Qy         89 TDGRTALDLA------DPSAKAVLTG---EYKKDE----LLESARSGNEEKMMALL--TP 133
Db        349 DSFESPLTLAACGGHVELAALLIERGASLEEVNDEGYTPLMEAAREGHEE-MVALLLGQG 407

Qy        134 LNVVNCHASDGRK----------------------------STPLHLAAGYNRVKIVQLL 164
Db        408 ANINAQTEETQETALTLACCGGFLEVADFLIKAGADIELGCSTPLMEAAQEGHLELVKYL 467

Qy        165 LQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRV 224
Db        468 LAAGANVHATTATGDTALTYACENGHTDVADVLLQAGADLEHESEGGRTPLMKAARAGHV 527

Qy        225 ---------------------------------EVCSLLLSYGADPT------------ 238
Db        528 CTVQFLISKGANVNRTTANNDHTVLSLACAGGHLAVVELLLAHGADPTHRLKDGSTMLIE 587

Qy        239 -------------------LLNCHNKSAIDLAP-------------------------- 252
Db        588 AAKGGHTSVVCYLLDYPNNLLSAPPPDVTQLTPPSHDLNRAPRVPVQALPMVVPPQEPDK 647

Qy        253 ----------------------------------TPQLKERLAYEFKGH--SLLQAAR 274
Db        648 PPANVATTLPIRNKAASKQKSSSHLPANSQDVQGYITNQSPESIVEEAQGKLTELEQRIK 707

Qy        275 EADVTRIKKHLSLEMVNFKHPQ--THE--TALHCAAASPYPKRKQICELLLRKGANINEK 330
Db        708 EA----IEKNAQLQSLELAHADQLTKEKIEELNKTREEQIQKKQKILEELQKVERELQLK 763

Qy        331 TKE---------------------------FLTPLHVASEKAHND------VVEVVVKH 356
Db        764 TQQQLKKQYLEVKAQRIQLQQQQQQSCQHLGLLTPVGVGEQLSEGDYARLQQVDPVLLKD 823

Qy        357 EAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQ--MGNENVQQL 414
Db        824 EPQQTAA-QMGFAPIQPLAMPQALPLAAGPLPPGSIANLTELQGVIVGQPVLGQAQLAGL 882

Qy        415 -------LQEGISLG------NSEADRQLLEAAKAGDVETVKKLCTVQSVN--------- 452
Db        883 GQGILTETQQGLMVASPAQTLNDTLDDIMAVSGRASAMSNTPTHSIAASISQPQTPTPSP 942

Qy        453 ------------CRDI----EGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 496
Db        943 IISPSAMLPIYPAIDIDAQTESNHDTALTLACAGGHEELVQTLLERGASIEHRDKKGFTP 1002

Qy        497 LHNACSYGHYEVAELLVKHGAVVNV-ADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKK 555
Db       1003 LILAATAGHVGVVEILLDNGADIEAQSERTKDTPLSLACSGGRQEVVELLLARGANKEHR 1062

Qy        556 NRDGNTPLDLVKDGD--TDIQDLLR------------GDAALLDAAKKGCLARVK-KLS 599
Db       1063 NVSDYTPLSLAASGGYVNIIKILLNAGAEINSRTGSKLGISPLMLAAMNGHTAAVKLLLD 1122

Qy        600 SPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHV 659
Db       1123 MGSDINAQIETNR-NTALTLACFQGRTEVVSLLLDRKANVEHRAKTGLTPLMEAASGGYA 1181

Qy        660 DVAALLIKYNACVNA--TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717
Db       1182 EVGRVLLDKGADVNAPPVPSSRDTALTIAADKGHYKFCELLIGRGAHIDVRNKKGNTPLW 1241

Qy        718 LVSAD---DVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSP 766
Db       1242 LAANGGHLDVVQLLVQA--------------------GADVDAADNRKITP 1272


RESULT 44
Q99NH0
ID   Q99NH0      PRELIMINARY;      PRT;  1599 AA.
AC   Q99NH0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Gene trap ankyrin repeat containing protein.
GN   GTAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RA   Watt A.J., Jones E.A., Ure J., Peddie D., Wilson D., Forrester L.M.;
RT   "A gene trap integration provides an early in situ marker for hepatic
RT   specification of the foregut endoderm.";
RL   Mech. Dev. 100:205-215(2001).
DR   EMBL; AY026253; AAK07672.1; -.
DR   HSSP; P42773; 1IHB.
DR   MGD; MGI:1932101; Gtar.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0003684; F:damaged DNA binding; IEA.
DR   GO; GO:0006298; P:mismatch repair; IEA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR000432; MutS_C.
DR   Pfam; PF00023; ank; 25.
DR   Pfam; PF00488; MutS_V; 1.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 25.
DR   PROSITE; PS50088; ANK_REPEAT; 20.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   ANK repeat; Repeat.
SQ   SEQUENCE   1599 AA;  170271 MW;  CCBB6C56D8ECB7C8 CRC64;

  Query Match             9.3%;  Score 534.5;  DB 11;  Length 1599;
  Best Local Similarity   24.7%;  Pred. No. 3.1e-25;
  Matches  257;  Conservative  116;  Mismatches  344;  Indels  325;  Gaps   30;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDN------ 56
Db        341 GYVDVVEVLLESGASIGDHNENGHTPLMEAGSAGHVEVARLLLENGAGINTHSNEFKESA 400

Qy         57 ----------------------------WNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRN 88
Db        401 LTLACYKGHLEMVRFLLEAGADQEHKTDEMHTALMEACMDGHVEVARLLLDSGAQVNMPA 460

Qy         89 TDGRTALDLA------DPSAKAVLTG---EYKKDE----LLESARSGNEEKMMALL--TP 133
Db        461 DSFESPLTLAACGGHVELAALLIERGASLEEVNDEGYTPLMEAAREGHEE-MVALLLGQG 519

Qy        134 LNVNCHASDGRK----------------------------STPLHLAAGYNRVKIVQLL 164
Db        520 ANINAQTEETQETALTLACCGGFLEVADFLIKAGADIELGCSTPLMEAAQEGHLELVKYL 579

Qy        165 LQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRV 224
Db        580 LAAGANVHATTATGDTALTYACENGHTDVADVLLQAGADLEHESEGGRTPLMKAARAGHV 639

Qy        225 ---------------------------------EVCSLLLSYGADPT------------ 238
Db        640 CTVQFLISKGANVNRTTANNDHTVLSLACAGGHLAVVELLLAHGADPTHRLKDGSTMLIE 699

Qy        239 ------------LLNCHN-----------------------------------------K 245
Db        700 AAKGGHTSVVCYLLDYPNNLLAAPPPDVTQLTPPSHDLNRAPRVPVQALPMVVPPQEPDK 759

Qy        246 SAIDLAPTPQLKERLAYEPKGHSLLQAAREADVTRIKKHLSLEMVNPKHPQTHETALHCA 305
Db        760 PPANLAATLPVRSKAASKQKSNSHL-PANSQDVQGYITNQSPESI-VEEAQGKLTELEQR 817

Qy        306 AASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVV-----VKHEAKV 360
Db        818 IKEAIEKNAQLQSLEL---AHADQLTKEKIEELNKTREEQIQKKQKILEELQKVERELQL 874

Qy        361 NALDNLGQTSLHRAAY-----------CGHLQTCRLLLSYG--------------CDPN 394
Db        875 KTQQQLKKQYLEVKAQRIQLQQQQQQSCQHL---GLLTSVGVGEQLSEGDYARLQQVDPV 931

Qy        395 IISLQ-GFTALQMGNENVQQLLQ----------------------EGISLGN-------- 423
Db        932 LLKDEPQQTAAQMGFAPIQPLAMPQALPLATGPLPPGSIANLTELQGVIVGQPVLGQAQL 991
```

```
Qy        424 --------SEADRQLLEAAKAGDVETVKKLCTVQSVNCR---------------------- 454
Db        992 AGLGQGILTETQQGLMVASPAQTLNDT--LDDIMAVSGRASAMSNTPTHSIAASVSQPQT 1049

Qy        455 -------------------DI----EGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDK 491
Db       1050 PTPSPIISPSAMLPIYPAIDIDAQTESNHDTALTLACAGGHEELVQTLLERGASIEHRDK 1109

Qy        492 GGLVPLHNACSYGHYEVAELLVKHGAVVNV-ADLWKFTPLHEAAAKGKYEICKLLLQHGA 550
Db       1110 KGFTPLILAATAGHVGVVEILLDNGADIEAQSERTKDTPLSLACSGGRQEVVELLLARGA 1169

Qy        551 DPTKKNRDGNTPLDLVKDGD--TDIQDLLR-------------GDAALLDAAKKGCLARV 595
Db       1170 NKEHRNVSDYTPLSLAASGGYVNIIKILLNAGAEINSRTGSKLGISPLMLAAMNGHTAAV 1229

Qy        596 K-KLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 654
Db       1230 KLLLDMGSDINAQIETNR-NTALTLACFQGRTEVVSLLLDRKANVEHRAKTGLTPLMEAA 1288

Qy        655 SYGHVDVAALLIKYNACVNA--TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
Db       1289 SGGYAEVGRVLLDKGADVNAPPVPSSRDTALTIAADKGHYKFCELLIGKGAHIDVRNKKG 1348

Qy        713 QTPLDLVSAD---DVSALLTAA 731
Db       1349 NTPLWLAANGGHLDVVQLLVQA 1370
```

```
RESULT 45
Q8TB46
ID   Q8TB46      PRELIMINARY;      PRT;   743 AA.
AC   Q8TB46;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein KIAA1223.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024725; AAH24725.1; -.
DR   InterPro; IPR002110; ANK.
DR   Pfam; PF00023; ank; 13.
DR   PRINTS; PR01415; ANKYRIN.
DR   SMART; SM00248; ANK; 12.
DR   PROSITE; PS50088; ANK_REPEAT; 13.
DR   PROSITE; PS50297; ANK_REP_REGION; 1.
KW   Hypothetical protein; ANK repeat; Repeat.
SQ   SEQUENCE   743 AA;  80197 MW;  FFFC6FD1264D172F CRC64;

  Query Match             9.1%;  Score 527;  DB 4;  Length 743;
  Best Local Similarity   24.0%;  Pred. No. 2.8e-25;
  Matches  201;  Conservative   98;  Mismatches  238;  Indels  302;  Gaps   19;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNA-----CSFGHAEVVNLLLRHGADPNARDNW 57
Db          2 GHREIVEHLLDHGAEVNHEDVDGRTALSVAALCVPASKGHASVVSLLIDRGAEVDHCDKD 61

Qy         58 NYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLE 117
Db         62 GMTPLLVAAYEGHVDVVDLLLEGGAD--VDHTD-------------------------- 92

Qy        118 SARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 177
Db         93 ----------------------NNGR--TPLLAAASMGHASVVNTLLFWGAAVDSIDSE 127

Qy        178 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 237
```

```
Db        128 GRTVLSIASAQGNVEVVRTLLDRGLDENHRDDAGWTPLHMAA------------------ 169

Qy        238 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 297

Db        170 --------------------------FEGHRL---------------------------- 175

Qy        298 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 357

Db        176 ------------------ICEALIEQGARTNEIDNDGRIPFILASQEGHYDCVQILLENK 217

Qy        358 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQE 417

Db        218 SNIDQRGYDGRNALRVAALEGHRDIVELLFSHGAD------------------------- 252

Qy        418 GISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVE 477

Db        253 --------------------------------VNCKDADGRPT--LYILALENQLTMAE 277

Qy        478 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 537

Db        278 YFLENGANVEASDAEGRTALHVSCWQGHMEMVQVLIAYHADVNAADNEKRSALQSAAWQG 337

Qy        538 KYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597

Db        338 HVKVVQLLIEHG--------------------------------AVVD------------ 353

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657

Db        354 ----------HTCNQGATALCIAAQEGHIDVVQVLLEHGADPNHADQFGRTAMRVAAKNG 403

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK-NQEGQTPL 716

Db        404 HSQIIKLLEKYGA--SSLNGCSPSPVHTMEQKPLQSLSSKV----QSLTIKSNSSGST-- 455

Qy        717 DLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSL--SAASS 774

Db        456 ---GGGD---------MQPS----------LRGL--PNGPTHAFSSPSESPDSTVDRQKSS 492

Qy        775 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQI 833

Db        493 LSNNSLKSSKNSSLRTTSSTATAQT-----VPIDSFHNLSFTEQIQQHSLPRSRSRQSI 546
```

Search completed: March 25, 2004, 16:43:54
Job time : 61.188 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:44:37 ; Search time 44.188 Seconds
(without alignments)
6513.709 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5766	99.9	1166	10	US-09-972-115A-6	Sequence 6, Appli
2	5766	99.9	1166	14	US-10-163-587A-15	Sequence 15, Appl
3	5766	99.9	1166	14	US-10-199-937-135	Sequence 135, App
4	5766	99.9	1169	14	US-10-199-937-2	Sequence 2, Appli
5	5766	99.9	1169	14	US-10-199-937-101	Sequence 101, App
6	5766	99.9	1262	14	US-10-199-937-107	Sequence 107, App
7	5766	99.9	1385	14	US-10-199-937-133	Sequence 133, App
8	5602	97.1	1074	9	US-09-509-196A-2	Sequence 2, Appli
9	5547.5	96.2	1227	10	US-09-849-602-26	Sequence 26, Appl
10	4923.5	85.3	1333	10	US-09-972-115A-2	Sequence 2, Appli
11	4889.5	84.8	1327	9	US-09-841-835-2	Sequence 2, Appli
12	4889.5	84.8	1327	10	US-09-972-115A-8	Sequence 8, Appli
13	4889.5	84.8	1327	14	US-10-199-937-4	Sequence 4, Appli
14	4847	84.0	1267	10	US-09-972-115A-4	Sequence 4, Appli
15	4102	71.1	784	14	US-10-199-937-89	Sequence 89, Appl
16	3959	68.6	756	14	US-10-199-937-91	Sequence 91, Appl
17	3835.5	66.5	1181	14	US-10-199-937-139	Sequence 139, App
18	3596	62.3	802	9	US-09-964-899-41	Sequence 41, Appl
19	3093	53.6	949	9	US-09-841-835-10	Sequence 10, Appl
20	2367	41.0	522	14	US-10-199-937-99	Sequence 99, Appl
21	2165	37.5	415	12	US-10-276-774-1690	Sequence 1690, Ap
22	2053.5	35.6	1099	14	US-10-199-937-178	Sequence 178, App
23	1918.5	33.3	673	9	US-09-841-835-8	Sequence 8, Appli
24	1895	32.8	362	14	US-10-199-937-160	Sequence 160, App
25	874	15.1	160	10	US-09-972-115A-20	Sequence 20, Appl
26	874	15.1	160	14	US-10-199-937-46	Sequence 46, Appl
27	863.5	15.0	1762	14	US-10-205-194-117	Sequence 117, App
28	860.5	14.9	1330	15	US-10-108-260A-3237	Sequence 3237, Ap
29	798.5	13.8	1724	9	US-09-964-899-43	Sequence 43, Appl
30	772	13.4	153	14	US-10-106-698-6500	Sequence 6500, Ap
31	707.5	12.3	1053	12	US-10-221-278-343	Sequence 343, App
32	707.5	12.3	1053	15	US-10-291-172-343	Sequence 343, App
33	678	11.8	919	15	US-10-108-260A-4122	Sequence 4122, Ap
34	650	11.3	119	14	US-10-199-937-42	Sequence 42, Appl
35	650	11.3	119	14	US-10-199-937-44	Sequence 44, Appl
36	639	11.1	118	14	US-10-199-937-8	Sequence 8, Appli
37	583	10.1	1619	14	US-10-369-978-4	Sequence 4, Appli
38	541	9.4	1704	14	US-10-369-978-2	Sequence 2, Appli
39	536.5	9.3	1709	14	US-10-369-978-6	Sequence 6, Appli
40	528.5	9.2	3913	15	US-10-334-143-45	Sequence 45, Appl
41	501.5	8.7	747	15	US-10-094-749-1924	Sequence 1924, Ap
42	474.5	8.2	657	15	US-10-104-047-2529	Sequence 2529, Ap
43	465	8.1	765	14	US-10-128-174-3	Sequence 3, Appli
44	465	8.1	765	14	US-10-128-174-34	Sequence 34, Appl
45	465	8.1	765	14	US-10-128-174-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-972-115A-6
; Sequence 6, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-6

Query Match 99.9%; Score 5766; DB 10; Length 1166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
```

```
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100

Db       1147 AYPEYLITYQIMRPEGMVDG 1166

RESULT 2
US-10-163-587A-15
; Sequence 15, Application US/10163587A
; Publication No. US20030096263A1
; GENERAL INFORMATION:
;  APPLICANT: Oliveira, Marcos
;  TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZING
;  FILE REFERENCE: 50229-306
;  CURRENT APPLICATION NUMBER: US/10/163,587A
;  CURRENT FILING DATE:  2003-01-10
;  PRIOR APPLICATION NUMBER: 60/296,110
;  PRIOR FILING DATE: 2001-06-07
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 1166
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-163-587A-15

  Query Match            99.9%;  Score 5766;  DB 14;  Length 1166;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
```

```
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166
```

RESULT 3
US-10-199-937-135
; Sequence 135, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-135

Query Match 99.9%; Score 5766; DB 14; Length 1166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166
```

RESULT 4
US-10-199-937-2
; Sequence 2, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29

```
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1169
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-2

  Query Match             99.9%;  Score 5766;  DB 14;  Length 1169;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         70 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 129

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 189

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 249

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 309

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 369

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 429

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 489

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 549

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 609

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 669

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 729

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 789

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 849

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 909

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 969

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1029

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1089

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1149

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1150 AYPEYLITYQIMRPEGMVDG 1169


RESULT 5
US-10-199-937-101
; Sequence 101, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
;   LENGTH: 1169
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-101

  Query Match             99.9%;  Score 5766;  DB 14;  Length 1169;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         70 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 129

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 189

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 249

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 309

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 369

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 429

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 489
```

```
Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 549

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 609

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 669

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 729

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 789

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 849

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 909

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 969

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1029

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1089

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1149

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1150 AYPEYLITYQIMRPEGMVDG 1169
```

```
RESULT 6
US-10-199-937-107
; Sequence 107, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
;   LENGTH: 1262
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-107
```

```
Query Match           99.9%;  Score 5766;  DB 14;  Length 1262;
Best Local Similarity  99.9%;  Pred. No. 0;
Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 222

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 282

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 342

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 402

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 462

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 522

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 582

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 642

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 702

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 762

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 822

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 882

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 942

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 1002

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1062

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1122

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1182
```

```
Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1242

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1243 AYPEYLITYQIMRPEGMVDG 1262


RESULT 7
US-10-199-937-133
; Sequence 133, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 133
;   LENGTH: 1385
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-133

  Query Match             99.9%;  Score 5766;  DB 14;  Length 1385;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        286 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 345

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 405

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 465

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 525

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 585

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 645

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 705

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 765

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 825

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 885

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 945

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 1005

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 1065

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 1125

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1185

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1245

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1305

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1365

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1366 AYPEYLITYQIMRPEGMVDG 1385


RESULT 8
US-09-509-196A-2
; Sequence 2, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
;  APPLICANT:  DALY, Roger J.
;  APPLICANT:  SUTHERLAND, Robert L.
;  TITLE OF INVENTION:  A Potential Effector for the Grb7 Family of Signalling
;  TITLE OF INVENTION:  Proteins
;  FILE REFERENCE:  1871-129
;  CURRENT APPLICATION NUMBER: US/09/509,196A
;  CURRENT FILING DATE:  2000-03-23
;  PRIOR APPLICATION NUMBER:  P09388
;  PRIOR FILING DATE:  1997-09-23
;  PRIOR APPLICATION NUMBER:  PCT AU98/00795
;  PRIOR FILING DATE:  1998-09-23
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1074
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-509-196A-2

  Query Match             97.1%;  Score 5602;  DB 9;  Length 1074;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1068;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         27 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 60
```

```
Qy          87 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 120

Qy         147 TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNA 206
               |||||||||||||||||||| ||||||||||||||||||||||||||||||||| ||||
Db         121 TPLHLAAGYNRVKIVQLLLQHGRDVHAKDKGDLVPLHNACSYGHYEVTELLVKHGGCVNA 180

Qy         207 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKG 266
               |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db         181 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCKNKSAIDLAPTPQLKERLAYEFKG 240

Qy         267 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 300

Qy         327 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 360

Qy         387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 420

Qy         447 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 506
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480

Qy         507 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 566
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 540

Qy         567 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 600

Qy         627 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEA 686
               |||||||||||||||||||||||||||||||||||||||||| :||||||||||||||||
Db         601 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFTPLHEA 660

Qy         687 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 746
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 720

Qy         747 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 806
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 780

Qy         807 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 866
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 840

Qy         867 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 926
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 900

Qy         927 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 986
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 960

Qy         987 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1046
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1020

Qy        1047 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
               ||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        1021 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEEAYPEYLITYQIMRPEGMVDG 1074
```

RESULT 9
US-09-849-602-26
; Sequence 26, Application US/09849602

; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-26

Query Match 96.2%; Score 5547.5; DB 10; Length 1227;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1070; Conservative 3; Mismatches 25; Indels 3; Gaps 2;

```
Qy           1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
               ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         129 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 188

Qy          61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         189 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 248

Qy         121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         249 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 308

Qy         181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         309 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 368

Qy         241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         369 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 428

Qy         301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         429 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 488

Qy         361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         489 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 548

Qy         421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         549 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 608

Qy         481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         609 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 668

Qy         541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         669 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 728

Qy         601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         729 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 788

Qy         661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         789 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 848

Qy         721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         849 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 908
```

```
Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 968

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1028

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1088

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK- 1019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db       1089 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGVQFTKT 1148

Qy       1020 DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE 1079
              |   :     ||     ||| |       :  ||||||||||||||||||||||||||||
Db       1149 DLVTFATAAALL--PGNLGKVFPAVQCNENGTSPPGHHSVTGRPSVNGLALAEYVIYRGE 1206

Qy       1080 QAYPEYLITYQIMRPEGMVDG 1100
              |||||||||||||||||||||
Db       1207 QAYPEYLITYQIMRPEGMVDG 1227


RESULT 10
US-09-972-115A-2
; Sequence 2, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1333
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: The 'Xaa' at location 1 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: The 'Xaa' at location 2 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: The 'Xaa' at location 3 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: The 'Xaa' at location 4 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: The 'Xaa' at location 5 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (6)..(6)
;   OTHER INFORMATION: The 'Xaa' at location 6 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: The 'Xaa' at location 7 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: The 'Xaa' at location 8 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: The 'Xaa' at location 9 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: The 'Xaa' at location 10 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (11)..(11)
;   OTHER INFORMATION: The 'Xaa' at location 11 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (12)..(12)
;   OTHER INFORMATION: The 'Xaa' at location 12 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: The 'Xaa' at location 13 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (14)..(14)
;   OTHER INFORMATION: The 'Xaa' at location 14 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (15)..(15)
;   OTHER INFORMATION: The 'Xaa' at location 15 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (16)..(16)
;   OTHER INFORMATION: The 'Xaa' at location 16 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (17)..(17)
;   OTHER INFORMATION: The 'Xaa' at location 17 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (18)..(18)
;   OTHER INFORMATION: The 'Xaa' at location 18 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (19)..(19)
;   OTHER INFORMATION: The 'Xaa' at location 19 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
```

; LOCATION: (20)..(20)
; OTHER INFORMATION: The 'Xaa' at location 20 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: The 'Xaa' at location 21 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: The 'Xaa' at location 22 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: The 'Xaa' at location 23 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: The 'Xaa' at location 24 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: The 'Xaa' at location 25 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (26)..(26)
; OTHER INFORMATION: The 'Xaa' at location 26 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: The 'Xaa' at location 27 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: The 'Xaa' at location 28 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: The 'Xaa' at location 29 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (30)..(30)
; OTHER INFORMATION: The 'Xaa' at location 30 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: The 'Xaa' at location 31 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: The 'Xaa' at location 32 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (33)..(33)
; OTHER INFORMATION: The 'Xaa' at location 33 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (34)..(34)
; OTHER INFORMATION: The 'Xaa' at location 34 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (35)..(35)
; OTHER INFORMATION: The 'Xaa' at location 35 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (36)..(36)
; OTHER INFORMATION: The 'Xaa' at location 36 stands for Lys, Asn, Arg, Ser, Thr, Ile,

Query Match 85.3%; Score 4923.5; DB 10; Length 1333;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 962; Conservative 7; Mismatches 131; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||:  :|||||||: |   :||||||| |||||||: :||||||||||||||||||
Db        225 GFGRKNXXKYLLQNGANXQXLYNGGLIPLHXACSFGHAKXINLLLRHGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        705 QHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||||||||||||||| ||||||| ||||| ||||| ||||||||||||||||||||
Db        765 ICKLLLQHGADPTKKNRDGNTLLDLVKDGXTDIQDXLRGDAVLLDAAKKGCLARVKKXXF 824

Qy        601 PDNVNCRDTQGRHSTPLHLA--AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658
              ||||||||||||||||||||
Db        825 PDNVNCRDTQGRHSTPLHLAGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 884

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        885 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 944

Qy        719 VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNL 778
                      ||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db        945 XXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNL 1004

Qy        779 SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERFQITLDVL 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1005 SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVL 1064

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1065 VEMGHKELKEIXINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 1124
```

```
Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK-------KLWERYTHRRKEVS 951
              ||||||||||||||||||||||||||||||||||||||||       :  ||||||||||||
Db       1125 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNXXXXXAKIRHEERYTHRRKEVS 1184

Qy        952 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1185 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1244

Qy       1012 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1245 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1304

Qy       1072 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              |||||||||||||||||||||||||||||
Db       1305 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1333
```

RESULT 11
US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-841-835-2

Query Match 84.8%; Score 4889.5; DB 9; Length 1327;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:||| ||||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||:| |||||||::|||||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||||||||||:|||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||:|||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::||  |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:|||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:||||||:||| |||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:|||||||||||||||||||:|||||||||:|||||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:|||||||:|||||||||||||||||:|||||||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||:|||||||||||||||||||||||||||||||||||||||||||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: |||  |||| |||:|:|||          : :| |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              :||:  :|:  :||:  |:|  || |:| :|:|:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:|||||||||||||||||||||: |||| ||||| :    ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :||||||||:|||||||:|||||||||:::|||| |||| ||: ||:||||||||||
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy        1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
               ||||||||||:|||||||||||||| |||||:||||||| |||||||| ||||||||
Db        1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy        1079 EQAYPEYLITYQIMRPE 1095
               |||||||||||||:||
Db        1298 EQAYPEYLITYQIMKPE 1314
```

RESULT 12
US-09-972-115A-8
; Sequence 8, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-8

Query Match 84.8%; Score 4889.5; DB 10; Length 1327;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||:||| ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||:| |||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||:|||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||:|||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:||||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:|||||||||||||||||||:|||||||||:|||||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:||||||||:|||||||||||||||||||:|||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||| ||||||||||||||||||||||||||||||||||||:|||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: ||| |||| |||:|:|||           : :| |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              :||:  :|:  :||:  |:|  || |:| :|:|:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:|||||||||||||||||||||: |||| |||||  :    ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :|||||||||:|||||||:|||||||:::|||| |||| ||: ||:|||||||||||
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              ||||||||||:|||||||||||||| |||||:|||||| |||||||| ||||||||
Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095
              |||||||||||||:||
Db       1298 EQAYPEYLITYQIMKPE 1314
```

RESULT 13
US-10-199-937-4
; Sequence 4, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-4

Query Match 84.8%; Score 4889.5; DB 14; Length 1327;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||:||| ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||:| |||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||:|||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404
```

```
Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||:||||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| ||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:||||||||||||||||||:|||||||:||||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:||||||||:|||||||||||||||||||:|||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||| ||||||||||||||||||||||||||||||||||:|||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: |||  |||| |||:|:|||          : :|  |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFERE-QITLDVL 838
              :||:  :|:  :||:  |:|  || |:|  :|:||:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:||||||||||||||||||||: |||| ||||| :   ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :||||||||||:||||||||:||||||||:::|||| |||| ||: ||:||||||||||
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              ||||||||||:|||||||||||||| |||||:||||||| |||||||| ||||||||
Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095
              |||||||||||||:||
Db       1298 EQAYPEYLITYQIMKPE 1314
```

RESULT 14
US-09-972-115A-4
; Sequence 4, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1267
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: The 'Xaa' at location 1 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (42)..(42)
; OTHER INFORMATION: The 'Xaa' at location 42 stands for Lys, Arg, Thr, or Met.
; NAME/KEY: misc_feature
; LOCATION: (98)..(98)
; OTHER INFORMATION: The 'Xaa' at location 98 stands for Lys, Arg, Thr, or Met.
; NAME/KEY: misc_feature
; LOCATION: (110)..(110)
; OTHER INFORMATION: The 'Xaa' at location 110 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (114)..(114)
; OTHER INFORMATION: The 'Xaa' at location 114 stands for Asn, Ser, Thr, or Ile.
; NAME/KEY: misc_feature
; LOCATION: (126)..(126)
; OTHER INFORMATION: The 'Xaa' at location 126 stands for Lys, or Asn.
; NAME/KEY: misc_feature
; LOCATION: (139)..(139)
; OTHER INFORMATION: The 'Xaa' at location 139 stands for Ile, Val, Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (144)..(144)
; OTHER INFORMATION: The 'Xaa' at location 144 stands for Thr.
; NAME/KEY: misc_feature
; LOCATION: (153)..(153)
; OTHER INFORMATION: The 'Xaa' at location 153 stands for Thr.
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: The 'Xaa' at location 166 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: The 'Xaa' at location 173 stands for Ile, Val, Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (174)..(174)
; OTHER INFORMATION: The 'Xaa' at location 174 stands for Glu, or Lys.
; NAME/KEY: misc_feature
; LOCATION: (183)..(183)
; OTHER INFORMATION: The 'Xaa' at location 183 stands for Val.
; NAME/KEY: misc_feature
; LOCATION: (185)..(185)
; OTHER INFORMATION: The 'Xaa' at location 185 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: The 'Xaa' at location 188 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (202)..(202)
; OTHER INFORMATION: The 'Xaa' at location 202 stands for His, or Pro.
; NAME/KEY: misc_feature
; LOCATION: (204)..(204)
; OTHER INFORMATION: The 'Xaa' at location 204 stands for Glu, or Lys.
; NAME/KEY: misc_feature
; LOCATION: (205)..(205)
; OTHER INFORMATION: The 'Xaa' at location 205 stands for Asn, Ser, Thr, or Ile.
; NAME/KEY: misc_feature
; LOCATION: (207)..(207)

; OTHER INFORMATION: The 'Xaa' at location 207 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (211)..(211)
; OTHER INFORMATION: The 'Xaa' at location 211 stands for Gln, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (213)..(213)
; OTHER INFORMATION: The 'Xaa' at location 213 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (215)..(215)
; OTHER INFORMATION: The 'Xaa' at location 215 stands for Lys, Asn, Glu, Asp, Gln, His
; OTHER INFORMATION: , a stop codon, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (227)..(227)
; OTHER INFORMATION: The 'Xaa' at location 227 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (228)..(228)
; OTHER INFORMATION: The 'Xaa' at location 228 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (237)..(237)
; OTHER INFORMATION: The 'Xaa' at location 237 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (238)..(238)
; OTHER INFORMATION: The 'Xaa' at location 238 stands for Asn, Ser, Thr, Ile, Asp, Gly
; OTHER INFORMATION: , Ala, Val, His, Arg, Pro, Leu, Tyr, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (667)..(667)
; OTHER INFORMATION: The 'Xaa' at location 667 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (725)..(725)
; OTHER INFORMATION: The 'Xaa' at location 725 stands for Lys, or Asn.
; NAME/KEY: misc_feature
; LOCATION: (727)..(727)
; OTHER INFORMATION: The 'Xaa' at location 727 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (735)..(735)
; OTHER INFORMATION: The 'Xaa' at location 735 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (741)..(741)
; OTHER INFORMATION: The 'Xaa' at location 741 stands for Met, Val, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (763)..(763)
; OTHER INFORMATION: The 'Xaa' at location 763 stands for Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (764)..(764)
; OTHER INFORMATION: The 'Xaa' at location 764 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (787)..(787)
; OTHER INFORMATION: The 'Xaa' at location 787 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (788)..(788)
; OTHER INFORMATION: The 'Xaa' at location 788 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (789)..(789)
; OTHER INFORMATION: The 'Xaa' at location 789 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (790)..(790)
; OTHER INFORMATION: The 'Xaa' at location 790 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (791)..(791)
; OTHER INFORMATION: The 'Xaa' at location 791 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (792)..(792)
; OTHER INFORMATION: The 'Xaa' at location 792 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (793)..(793)
; OTHER INFORMATION: The 'Xaa' at location 793 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (794)..(794)
; OTHER INFORMATION: The 'Xaa' at location 794 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (795)..(795)
; OTHER INFORMATION: The 'Xaa' at location 795 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (796)..(796)
; OTHER INFORMATION: The 'Xaa' at location 796 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: The 'Xaa' at location 797 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (798)..(798)
; OTHER INFORMATION: The 'Xaa' at location 798 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (799)..(799)
; OTHER INFORMATION: The 'Xaa' at location 799 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (800)..(800)
; OTHER INFORMATION: The 'Xaa' at location 800 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (801)..(801)
; OTHER INFORMATION: The 'Xaa' at location 801 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (802)..(802)

Query Match 84.0%; Score 4847; DB 10; Length 1267;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 949; Conservative 6; Mismatches 143; Indels 4; Gaps 2;

```
Qy          2 FGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 61
              |||||:  ||| |||: | || |||||||||||||| |  : ||| | | ||||||||||
Db        167 FGRKDLXXYLLTNGANXQXRDXGGLIPLHNACSFGXAXXIXLLLXHXAXPNARDNWNYTP 226

Qy         62 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 121
              ||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 XXEAAIKGKIXXCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 286

Qy        122 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 346

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 406

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 466
```

```
Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 526

Qy        362 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 586

Qy        422 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 646

Qy        482 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 541
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        647 HGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 706

Qy        542 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 601
              |||||||||||||| : |   ||||||| ||||| ||||| |||||||||||||||   |
Db        707 CKLLLQHGADPTKKKKKGXIXLDLVKDGXTDIQDXLRGDAVLLDAAKKGCLARVKKXXFP 766

Qy        602 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 661
              |||||||||||||||| ||
Db        767 DNVNCRDTQGRHSTPLHL-AGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 825

Qy        662 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 721

Db        826 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 885

Qy        722 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 781
                      ||||||| |||| ||||||||||||||||||||||||||||||||||||||||||||
Db        886 XXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 945

Qy        782 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFERFQITLDVLVEM 841
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVLVEM 1005

Qy        842 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 901
              |||||||| ||||||||||   ||||||||||||||||||||||||||||||||||||||||
Db       1006 GHKELKEIXINAYGHRHKLIKSFERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1065

Qy        902 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK---KLWERYTHRRKEVSEENHNHA 958
              |||||||||||||||||||||||||||||||:   :  ||||||||||||||||||||
Db       1066 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNRAKIRHEERYTHRRKEVSEENHNHA 1125

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1185

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1245

Qy       1079 EQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||
Db       1246 EQAYPEYLITYQIMRPEGMVDG 1267


RESULT 15
US-10-199-937-89
; Sequence 89, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
;   LENGTH: 784
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-89

  Query Match             71.1%;  Score 4102;  DB 14;  Length 784;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  784;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        317 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 60

Qy        377 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 120

Qy        437 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 180

Qy        497 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 240

Qy        557 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 300

Qy        617 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 360

Qy        677 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 420

Qy        737 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 480

Qy        797 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 540

Qy        857 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 600

Qy        917 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 660

Qy        977 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 720

Qy       1037 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 780

Qy       1097 MVDG 1100
              ||||
Db        781 MVDG 784


RESULT 16
US-10-199-937-91
```

; Sequence 91, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-91

Query Match 68.6%; Score 3959; DB 14; Length 756;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        345 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 404
              ||||||
```

```
Db        599 YDICKLLLKHGADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRL 658

Qy        599 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658

Db        659 VTPESINCRDAQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGH 718

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718

Db        719 LDIAALLIKHKTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIEL 778

Qy        719 VSADDVSALLTAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAA 772

Db        779 ATADDVKCLLQDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPT 837

Qy        773 SSLDNLSGSFSELSSV-----VSSS-----------GTEGASSLEKKEVPGVD--FSIT 813

Db        838 TETVLLPTGASMILSVPVPLPLSSSTRISPAQGAEANGAEGSSSDDL--LPDADTITNVS 895

Qy        814 QFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 873

Db        896 GFLSSQQLHHLIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS---- 951

Qy        874 LNPYLTLNTSGSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYN 927

Db        952 --------TTGIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYN 1003

Qy        928 ILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGG 987

Db       1004 IIRVQKVQNRKLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGG 1063

Qy        988 MFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAM 1047

Db       1064 MFGAGIYFAEHSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAM 1123

Qy       1048 KMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100

Db       1124 KMAHAPPGHHSVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176
```

RESULT 18
US-09-964-899-41
; Sequence 41, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-41

Query Match 62.3%; Score 3596; DB 9; Length 802;
Best Local Similarity 65.1%; Pred. No. 4.1e-293;
Matches 720; Conservative 5; Mismatches 5; Indels 376; Gaps 5;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGE-----YKKDEL 115

Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGKSVYSGYSRKPV 186

Qy        116 LES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174

Db        187 KNNLARSGNEEKMMALLTPLNVNCHASDGR---------------------------- 216

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234

Db        217 -----------------------KHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 252

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294

Db        253 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 312

Qy        295 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 354

Db        313 PQTHETAL---------------------------------------------------- 320

Qy        355 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 414

Db        321 ---------------------------------------------------------- 320

Qy        415 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 474

Db        321 ----------------------------KLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 351

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 534

Db        352 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 411

Qy        535 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 594

Db        412 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 471

Qy        595 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 654

Db        472 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 531

Qy        655 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 714

Db        532 SYG---------------------------------------------------------- 534

Qy        715 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 774

Db        535 ---------------------------------------------------------- 534

Qy        775 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 834

Db        535 ----------------------------------------------------------IT 536

Qy        835 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 894

Db        537 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 596

Qy        895 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 954

Db        597 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 656

Qy        955 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1014

Db        657 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 716

Qy       1015 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1074

Db        717 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 776

Qy       1075 IYRGEQAYPEYLITYQIMRPEGMVDG 1100

Db        777 IYRGEQAYPEYLITYQIMRPEGMVDG 802
```

RESULT 19
US-09-841-835-10

```
; Sequence 10, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 949 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-841-835-10

  Query Match           53.6%;  Score 3093;  DB 9;  Length 949;
  Best Local Similarity  81.5%;  Pred. No. 1.2e-250;
  Matches  585;  Conservative   60;  Mismatches   71;  Indels    2;  Gaps    1;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||||:||| ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||:| ||||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||:||||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||:|||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |

Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||||||:||| ||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:|||||||||||||||||||:|||||||||:||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH-- 658
              |:|:|||||||:||||||||||||||||||:|||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCL 884

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 716
              | |  |  |  |||| ||  ::  || ||||  ::: |  ||
Db        885 ARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942


RESULT 20
US-10-199-937-99
; Sequence 99, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 99
;   LENGTH: 522
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-199-937-99

  Query Match           41.0%;  Score 2367;  DB 14;  Length 522;
  Best Local Similarity  99.8%;  Pred. No. 5.5e-190;
  Matches  452;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db         70 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 129

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 189

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 249

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 309

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 369

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 429

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 489

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 453
              |||||||||||||||||||||||||||||||||
Db        490 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 522


RESULT 21
US-10-276-774-1690
; Sequence 1690, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT:  Tang, Y, Tom et al
;  TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-030
;  CURRENT APPLICATION NUMBER: US/10/276,774
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 2700
;  SOFTWARE: Custom
; SEQ ID NO 1690
;   LENGTH: 415
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-276-774-1690

  Query Match             37.5%;  Score 2165;  DB 12;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-173;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        208 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 60

Qy        268 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 120

Qy        328 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 180

Qy        388 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 240

Qy        448 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 300

Qy        508 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 360

Qy        568 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 413


RESULT 22
US-10-199-937-178
; Sequence 178, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
;   LENGTH: 1099
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Parpla-Tank2b
;   OTHER INFORMATION: Fusion
US-10-199-937-178

  Query Match             35.6%;  Score 2053.5;  DB 14;  Length 1099;
  Best Local Similarity   62.1%;  Pred. No. 4.7e-163;
  Matches  459;  Conservative   38;  Mismatches  115;  Indels  127;  Gaps   17;

Qy        404 LQMGNENVQQLLQEGISLG---NSEADRQLLEAAKAGDVETV-KKLCTVQSVNCRDIEGR 459
              |    : |: ::::   ||      |::  |  :   :|| : ||:  |:  | |
Db        437 LSRNKDEVKAMIEK---LGGKLTGTANKASLCISTKKEVEKMNKKMEEVKEANIR----- 488

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
                          | |  ||  || |  |         |    :| :|   || |
Db        489 -------------VVSEDFLQ---DVSASTK----------SLQELFLAHILSPWGAEV 521

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 579
                    |  |:   | :||          ||  :||:: |    : :   :  ::  |:|
Db        522 ------KAEPVEVVAPRGK---------SGAALSKKSK-GQVKEEGINKSEKRMKLTLKG 565

Qy        580 DAAL-----------------LDAAKKGCLARVK--------KLSSPDNVN---CRDTQ 610
              ||:                : :|  | :  ||        :|   |  |       :
Db        566 GAAVDPDSGLEHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSW 625

Qy        611 GRHSTPLHLAAGYNNLE---VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 667
              ||  | :    | | ||    |  ::|   :  :  |      |  |
Db        626 GRVGTVI----GSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSKNFTKYPK--------- 672

Qy        668 YNACVNATDKWAFTPLH------EAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 721
                        | ||       | | |  |           :|  |  ||||||||||
Db        673 -----------KFYPLEIDYGQDEEAVKKLT----------VNPGTKRVEGQTPLDLVSA 711

Qy        722 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 781
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 771

Qy        782 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 841
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 831

Qy        842 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 901
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 891

Qy        902 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 961
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 951

Qy        962 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1021
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1011

Qy       1022 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1081
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1012 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1071

Qy       1082 YPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1072 YPEYLITYQIMRPEGMVDG 1090
```

RESULT 23
US-09-841-835-8
; Sequence 8, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 673 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-841-835-8

Query Match 33.3%; Score 1918.5; DB 9; Length 673;
Best Local Similarity 80.9%; Pred. No. 4.7e-152;
Matches 364; Conservative 42; Mismatches 43; Indels 1; Gaps 1;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||:||| ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||:| ||||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||:||||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||:|||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
```

```
              ||| |||:|:||||:|:||| ||||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              |||||  || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||||| ||:|||||||||| ||||| |||:|
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYG 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQS 450
                  |    |   ||: |: | |::: :  |
Db        645 SDPSIISLQGFTAAQMGN-EAVQQILSGHS 673
```

RESULT 24
US-10-199-937-160
; Sequence 160, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-199-937-160

Query Match 32.8%; Score 1895; DB 14; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.6e-150;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        212 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 60

Qy        272 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 120

Qy        332 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 180

Qy        392 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 240

Qy        452 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 300

Qy        512 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 360

Qy        572 DI 573
              ||
Db        361 DI 362
```

```
RESULT 25
US-09-972-115A-20
; Sequence 20, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-20

  Query Match             15.1%;  Score 874;  DB 10;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-65;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        941 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 60

Qy       1001 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 120

Qy       1061 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||||||||||||||||||||
Db        121 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 160


RESULT 26
US-10-199-937-46
; Sequence 46, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
US-10-199-937-46

  Query Match             15.1%;  Score 874;  DB 14;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-65;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        941 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 60

Qy       1001 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 120

Qy       1061 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||||||||||||||||||||
Db        121 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 160


RESULT 27
US-10-205-194-117
; Sequence 117, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
;  APPLICANT: Warner-Lambert Company
;  APPLICANT:  Lee, Kevin
;  APPLICANT:  Dixon, Alistair
;  APPLICANT:  Brooksbank, Robert
;  APPLICANT:  Pinnock, Robert
;  TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
;  FILE REFERENCE: WL-A-018201
;  CURRENT APPLICATION NUMBER: US/10/205,194
;  CURRENT FILING DATE:  5200-07-24
;  PRIOR APPLICATION NUMBER: GB 0118354.0
;  PRIOR FILING DATE: 2001-07-27
;  NUMBER OF SEQ ID NOS: 177
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
;   LENGTH: 1762
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Ankyrin isoform
US-10-205-194-117

  Query Match             15.0%;  Score 863.5;  DB 14;  Length 1762;
  Best Local Similarity   30.2%;  Pred. No. 1.5e-62;
  Matches  270;  Conservative  123;  Mismatches  369;  Indels  133;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:|      |   || |   | |||| :|: :||: ||:    :|||
Db         79 GHVEVVSELLQREANVDQPTKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 138

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :||  ::   || | | :|     |     :|  : |   :
Db        139 YMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 198

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |  |   :: |  |||   : |          ||||:|| |  : :  |||   | |
Db        199 ALHIAARKDDTKAAALLLQNDTNADIESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 258

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |   |:  : :||:  || ::|      ||||  |     :|  :||
Db        259 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRA 318

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |   |                 | | |  | :|          ||:  :  :|
Db        319 A-PILSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 350

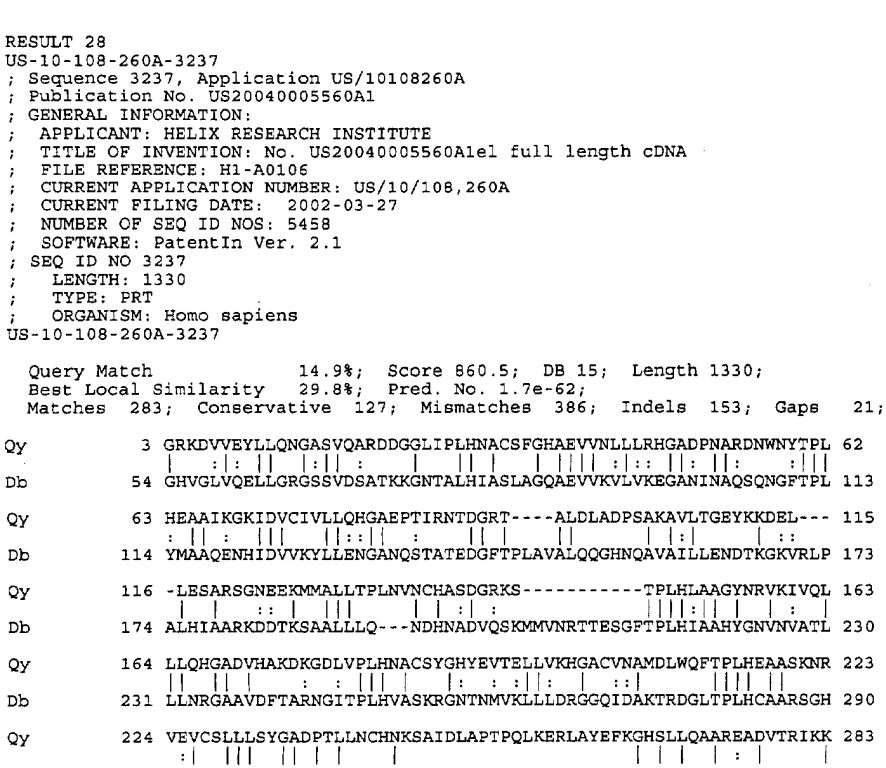
Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | || | |       ||||:| :|
Db        351 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTPLHIACKKNRI 407

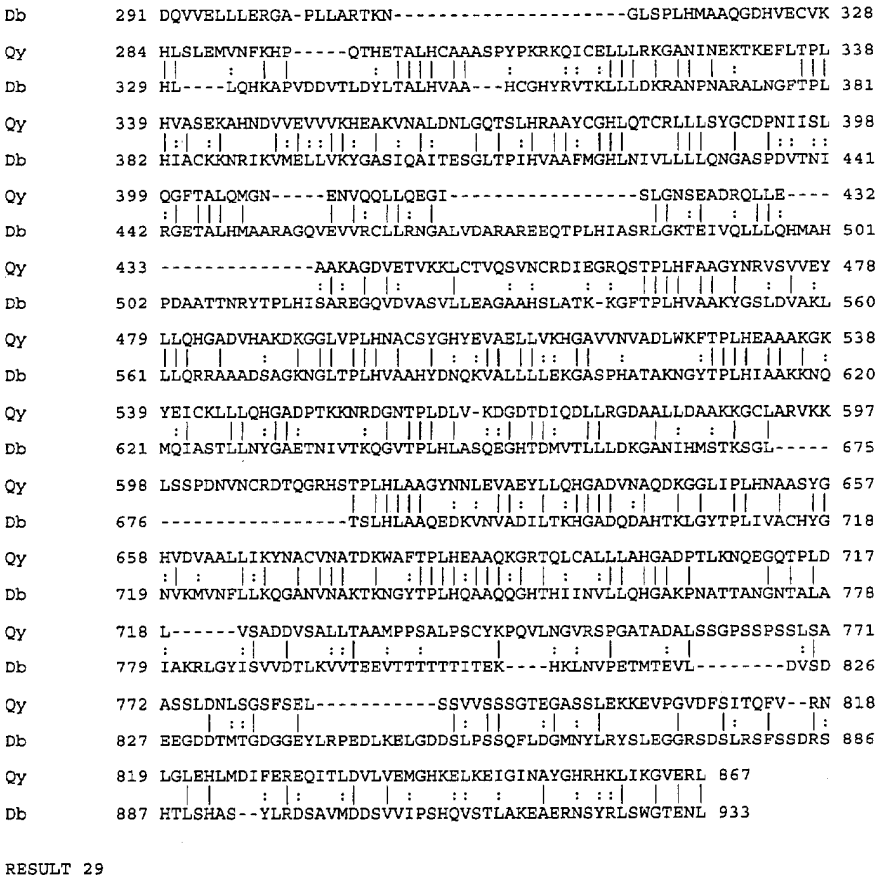
Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
              |:|:::|| | : |:   | | :| ||: ||:     |: :|  ||  :::| ||| |
Db        408 RVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRGETALHMA 467
```



RESULT 28
US-10-108-260A-3237
; Sequence 3237, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3237
; LENGTH: 1330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3237

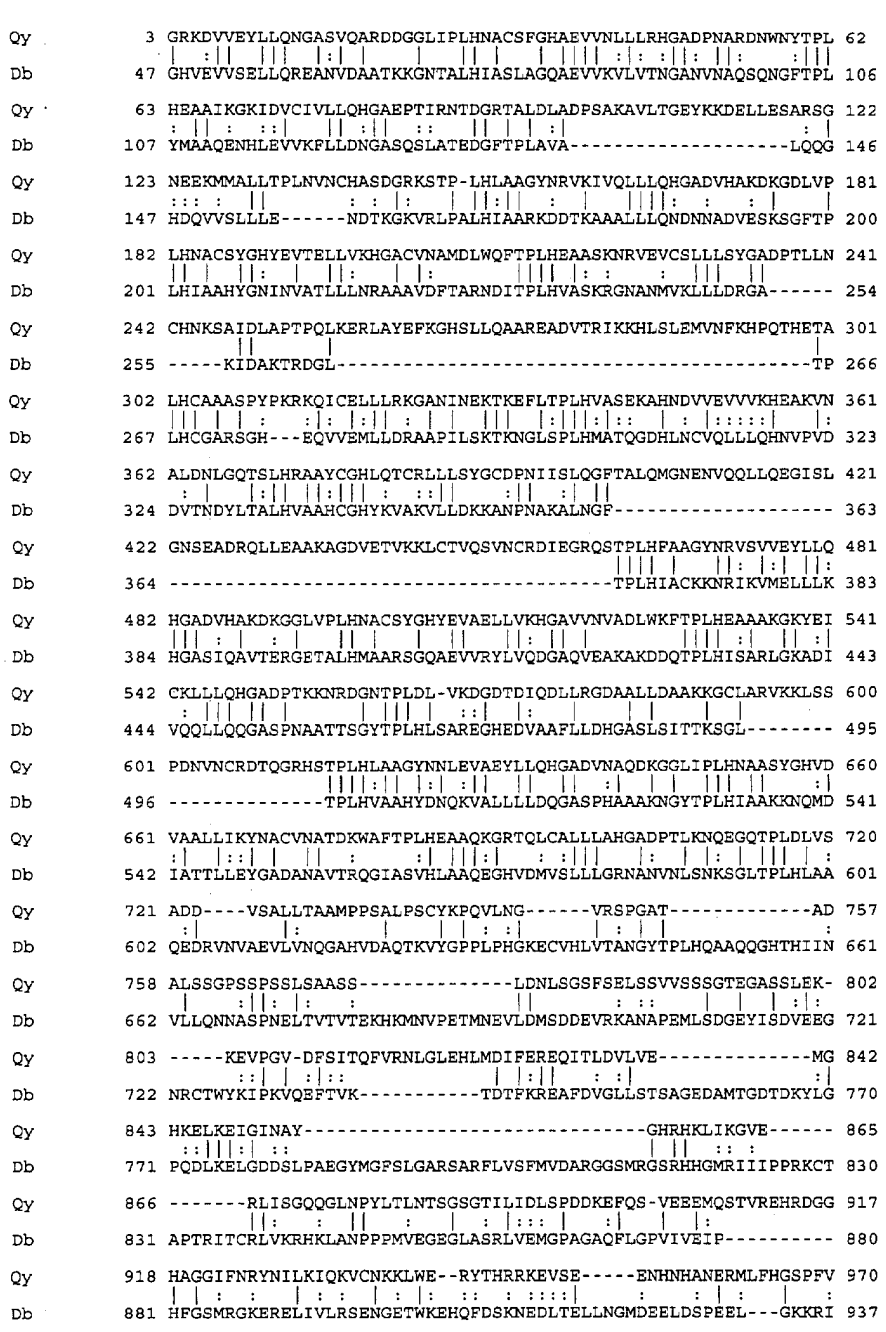
Query Match 14.9%; Score 860.5; DB 15; Length 1330;
Best Local Similarity 29.8%; Pred. No. 1.7e-62;
Matches 283; Conservative 127; Mismatches 386; Indels 153; Gaps 21;

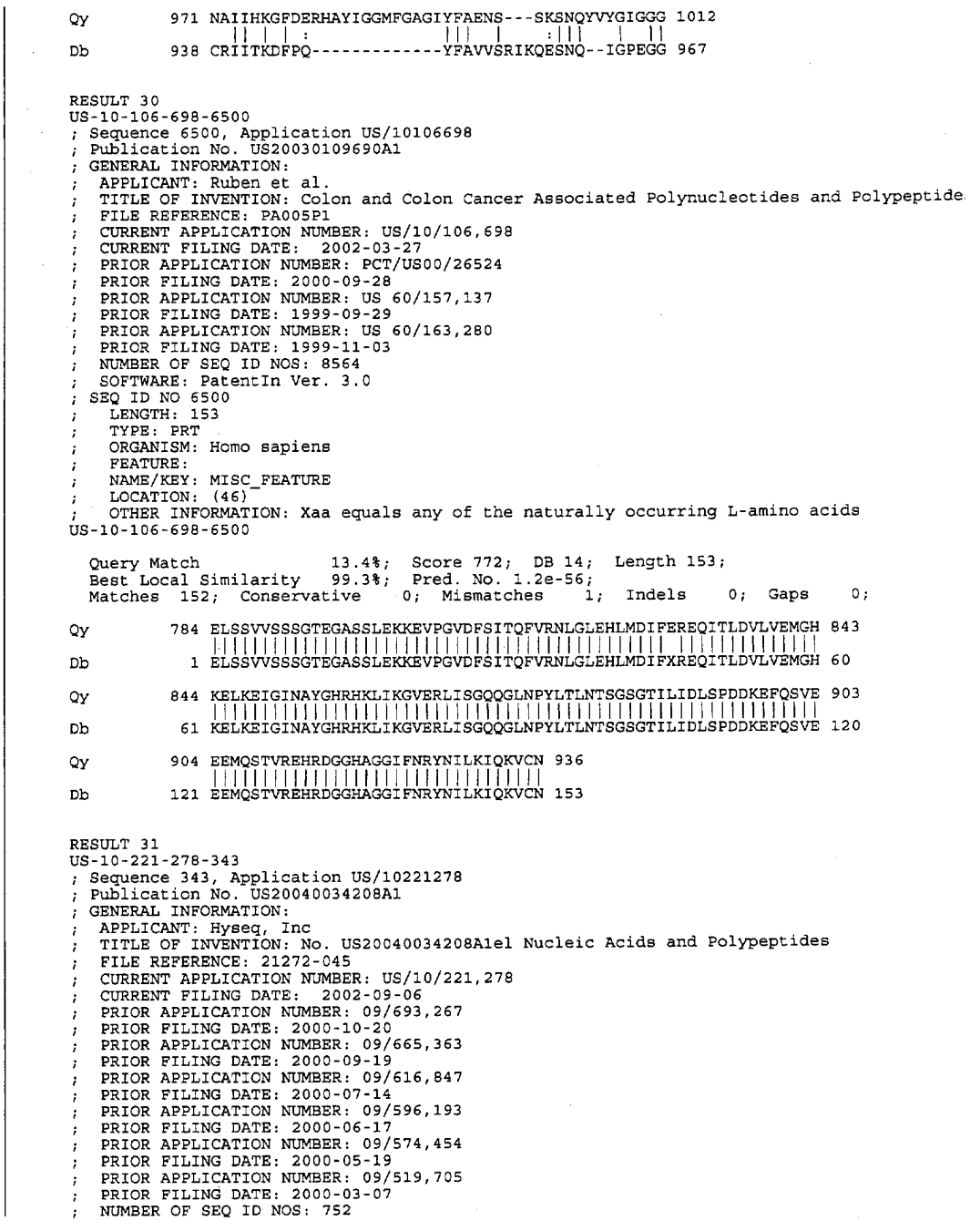




RESULT 29
US-09-964-899-43
; Sequence 43, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-964-899-43

Query Match 13.8%; Score 798.5; DB 9; Length 1724;
Best Local Similarity 25.1%; Pred. No. 4.3e-57;
Matches 282; Conservative 147; Mismatches 377; Indels 319; Gaps 27;





RESULT 30
US-10-106-698-6500
; Sequence 6500, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6500
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6500

Query Match 13.4%; Score 772; DB 14; Length 153;
Best Local Similarity 99.3%; Pred. No. 1.2e-56;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 31
US-10-221-278-343
; Sequence 343, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752

; SEQ ID NO 343
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-343

Query Match 12.3%; Score 707.5; DB 12; Length 1053;
Best Local Similarity 27.2%; Pred. No. 8.5e-50;
Matches 259; Conservative 130; Mismatches 367; Indels 197; Gaps 29;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         19 GDPDEVRALIFKKEDVNFQDNEKRTPLHAAAYLGDAEIIELLILSGARVNAKDSKWLTPL 78

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
Db         79 HRAVASCSEEAVQVLLKHSADVNARDKNWQTPLHIA-AANKAV----------------- 120

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPL 182
Db        121 ---KCAEALVPLLSNVNVSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFDKKDRRAI 177

Qy        183 HNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNC 242
Db        178 HWAAYMGHIEVVKLLVSHGAEVTCKDKKSYTPLHAAASSGMISVVKYLLDLGVDMNEPNA 237

Qy        243 HNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETA 301
Db        238 YGNTPLHVAC-----------YNGQDVV----------VNELIDCGAIVNQKNEKGF-TP 275

Qy        302 LHCAAASPYPKRKQIC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
Db        276 LHFAAASTH---GALCLELLVGNGADVNMKSKDGKTPLHMTALHGRFSRSQTIIQSGAVI 332

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQQLL 415
Db        333 DCEDKNGNTPLHIAARYGHELLINTLITSGADTAKRGIHGMFPLHLAALSGFSDCCRKLL 392

Qy        416 QEGISLGN-SEADRQLLEAAKA-GDVETVKKL----------CTVQ 449
Db        393 SSGFDIDTPDDFGRTCLHAAAAGGNLECLNLLLNTGADFNKKDKFGRSPLHYAAANCNYQ 452

Qy        450 ----------SVNCRDIEGRQSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGGLVPLH 498
Db        453 CLFALVGSGASVN--DLDERGCTPLHYAATSDTDGKCLEYLLRNDANPGIRDKQGYNAVH 510

Qy        499 NACSYGHYEVAELLVKHGAV-----VNVADLWK-------FTPLHEAAAKGKYEICKLLL 546
Db        511 YSAAYGHRLCLQLIASETPLDVLMETSGTDMLSDSDNRATISPLHLAAYHGHHQALEVLV 570

Qy        547 QHGADPTKKNRDGNTPLDLVK-DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVN 605
Db        571 QSLLDLDVRNSSGRTPLDLAAFKGHVECVDVLINQGASI-LVKDYILKR---------- 618

Qy        606 CRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNA---QDKGGLIPLHNAASYGHVDVA 662
Db        619 ---------TPIHAAATNGHSECLRLLIGNAEPQNAVDIQDGNGQTPLMLSVLNGHTDCV 669

Qy        663 ALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA- 721
Db        670 YSLLNKGANVDAKDKWGRTALHRGAVTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAAC 729

Qy        722 ---DDVSALLTAAMPPSALP-------------SCYKPQVLNG--------------VRS 751
Db        730 GHIGVLGALLQSAASMDANPATADNHGYTALHWACY-----NGHETCVELLLEQEVFQKT 784

Qy        752 PGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSL---------- 800
Db        785 EGNAFSPLHCAVINDNEGAAEMLIDTLG------ASIVNATDSKGRTPLHAAAFTDHVEC 838

Qy        801 ------EKKEVPGVDFSITQFVRNLGLEHLMDIFEREQI-TLDVLVEMGHKEL 846
Db        839 LQLLLSHNAQVNSVD--------STGKTPLMMAAENGQTNTVEMLVSSASAEL 883
```

RESULT 32
US-10-291-172-343
; Sequence 343, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 343
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-343

Query Match 12.3%; Score 707.5; DB 15; Length 1053;
Best Local Similarity 27.2%; Pred. No. 8.5e-50;
Matches 259; Conservative 130; Mismatches 367; Indels 197; Gaps 29;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         19 GDPDEVRALIFKKEDVNFQDNEKRTPLHAAAYLGDAEIIELLILSGARVNAKDSKWLTPL 78

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
Db         79 HRAVASCSEEAVQVLLKHSADVNARDKNWQTPLHIA-AANKAV----------------- 120

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPL 182
Db        121 ---KCAEALVPLLSNVNVSDRAGRTALHHAAFSGHGEMVKLLLSRGANINAFDKKDRRAI 177

Qy        183 HNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNC 242
Db        178 HWAAYMGHIEVVKLLVSHGAEVTCKDKKSYTPLHAAASSGMISVVKYLLDLGVDMNEPNA 237

Qy        243 HNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL-EMVNFKHPQTHETA 301
Db        238 YGNTPLHVAC-----------YNGQDVV----------VNELIDCGAIVNQKNEKGF-TP 275

Qy        302 LHCAAASPYPKRKQIC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
Db        276 LHFAAASTH---GALCLELLVGNGADVNMKSKDGKTPLHMTALHGRFSRSQTIIQSGAVI 332

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG-----NENVQQLL 415
Db        333 DCEDKNGNTPLHIAARYGHELLINTLITSGADTAKRGIHGMFPLHLAALSGFSDCCRKLL 392

Qy        416 QEGISLGN-SEADRQLLEAAKA-GDVETVKKL----------CTVQ 449
Db        393 SSGFDIDTPDDFGRTCLHAAAAGGNLECLNLLLNTGADFNKKDKFGRSPLHYAAANCNYQ 452

Qy        450 ----------SVNCRDIEGRQSTPLHFAAGYNR-VSVVEYLLQHGADVHAKDKGGLVPLH 498
Db        453 CLFALVGSGASVN--DLDERGCTPLHYAATSDTDGKCLEYLLRNDANPGIRDKQGYNAVH 510

Qy        499 NACSYGHYEVAELLVKHGAV-----VNVADLWK-------FTPLHEAAAKGKYEICKLLL 546
```

```
Db        511 YSAAYGHRLCLQLIASETPLDVLMETSGTDMLSDSDNRATISPLHLAAYHGHHQALEVLV 570

Qy        547 QHGADPTKKNRDGNTPLDLVK-DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVN 605

Db        571 QSLLDLDVRNSSGRTPLDLAAFKGHVECVDVLINQGASI-LVKDYILKR----------- 618

Qy        606 CRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNA---QDKGGLIPLHNAASYGHVDVA 662

Db        619 ---------TPIHAAATNGHSECLRLLIGNAEPQNAVDIQDGNGQTPLMLSVLNGHTDCV 669

Qy        663 ALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA- 721

Db        670 YSLLNKGANVDAKDKWGRTALHRGAVTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAAC 729

Qy        722 ---DDVSALLTAAMPPSALP-------------SCYKPQVLNG-------------VRS 751

Db        730 GHIGVLGALLQSAASMDANPATADNHGYTALHWACY-----NGHETCVELLLEQEVFQKT 784

Qy        752 PGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSL---------- 800

Db        785 EGNAFSPLHCAVINDNEGAAEMLIDTLG------ASIVNATDSKGRTPLHAAAFTDHVEC 838

Qy        801 ------EKKEVPGVDFSITQFVRNLGLEHLMDIFEREQI-TLDVLVEMGHKEL 846

Db        839 LQLLLSHNAQVNSVD--------STGKTPLMMAAENGQTNTVEMLVSSASAEL 883


RESULT 33
US-10-108-260A-4122
; Sequence 4122, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4122
;   LENGTH: 919
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-108-260A-4122

  Query Match           11.8%;  Score 678;  DB 15;  Length 919;
  Best Local Similarity   28.5%;  Pred. No. 2.1e-47;
  Matches  235;  Conservative  109;  Mismatches  304;  Indels  178;  Gaps   24;

Qy         11 LLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGK 70

Db          2 LIHKTEDVNTLDSEKRTPLHVAAFLGDAEIIELLILSGARVNAKDNMWLTPLHRAVASRS 61

Qy         71 IDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMAL 130

Db         62 EEAVQVLIKHSADVNARDKNWQTPLHVA-AANKAV-------------------KCAEV 100

Qy        131 LTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGH 190

Db        101 IIPLLSSVNVSDRGGRTALHHAALNGHVEMVNLLLAKGANINAFDKKDRRALHWAAYMGH 160

Qy        191 YEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDL 250

Db        161 LDVVALLINHGAEVTCKDKKGYTPLHAAASNGQINVVKHLLNLGVEIDEINVYGNTALHI 220

Qy        251 APTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHE-TALHCAAASP 309

Db        221 A---------CYN---------GQDAVVNELIDY----GANVNQPNNNGFTPLHFAAAST 258

Qy        310 YPKRKQIC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQ 368

Db        259 H---GALCLELLVNNGADVNIQSKDGKSPLHMTAVHGRFTRSQTLIQNGGEIDCVDKDGN 315

Qy        369 TSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADR 428

Db        316 TPLHVAARYGHELLINTLRTSGADTAKCGIHSMFPLHLAALN----------AHSDCCR 364

Qy        429 QLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHA 488

Db        365 KLL--SSGFEIDTPDKF-----------GR--TCLHAAAAGGNVECIKLLQSSGADFHK 408

Qy        489 KDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAA------------ 535

Db        409 KDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDDWGRTALHYAAASDMDRNKTILGNA 468

Qy        536 --------------KGKYEIC-KLLLQHGADPTKKNRDG---------------------- 559

Db        469 HDNSEELERARELKEKEATLCLEFLLQNDANPSIRDKEGYNSIHYAAAYGHRQCLELLLE 528

Qy        560 --------------NTPLDLVK-DG-------------DTDIQDLLRGDAALLDAAKKGC 591

Db        529 RTNSGFEESDSGATKSPLHLAAYNGHHQALEVLLQSLVDLDIRD-EKGRTALDLAAFKGH 587

Qy        592 LARVKKLSSP-------DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGAD----VN 640

Db        588 TECVEALINPGASIFVKDNVTKR-------TPLH-ASVINGHTLCLRLLLEIADNPEAVD 639

Qy        641 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 700

Db        640 VKDAKGQTPLMLAVAYGHIDAVSLLLEKEANVDTVDILGCTALHRGIMTGHEECVQMLLE 699

Qy        701 HGADPTLKNQEGQTPLDLVSADD----VSALLTAAMPPSALPSCYK 742

Db        700 QEVSILCKDSRGRTPLHYAAARGHATWLSELLQMAL--SEEDCCFK 743


RESULT 34
US-10-199-937-42
; Sequence 42, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 119
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
US-10-199-937-42

  Query Match           11.3%;  Score 650;  DB 14;  Length 119;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-46;
  Matches  119;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        962 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1021

Db          1 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 60

Qy       1022 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Db         61 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 119
```

RESULT 35
US-10-199-937-44
; Sequence 44, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence not
; OTHER INFORMATION: specified as protein-coding is vector sequence
US-10-199-937-44

Query Match 11.3%; Score 650; DB 14; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.5e-46;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 962 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1021
Db 1 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 60

Qy 1022 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
Db 61 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 119

RESULT 36
US-10-199-937-8
; Sequence 8, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa= unknown
US-10-199-937-8

Query Match 11.1%; Score 639; DB 14; Length 118;
Best Local Similarity 99.2%; Pred. No. 1.2e-45;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 963 LFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS 1022
Db 1 LFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS 60

Qy 1023 CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
Db 61 CYICHRXLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 118

RESULT 37
US-10-369-978-4
; Sequence 4, Application US/10369978
; Publication No. US20030152991A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/10/369,978
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US/09/392,812A
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1619
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC cDNA
; OTHER INFORMATION: sequence
US-10-369-978-4

Query Match 10.1%; Score 583; DB 14; Length 1619;
Best Local Similarity 26.8%; Pred. No. 5.3e-39;
Matches 236; Conservative 105; Mismatches 319; Indels 222; Gaps 26;

Qy 4 RKDV--VEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 61
Db 278 RRDVHMVRILVDYGTNVDTQNGEGQTPLHIAAAEGDEALLKYFYGVRASASIADNQDRTP 337

Qy 62 LHEAAIKGKIDVCIVLL-QHGAEPTIRNTDGRTALDLADPSAKA-VLTGEYKKDELLESA 119
Db 338 MHLAAENGHAHVIEILADKFKASIFERTKDGSTLMHIASLNGHAECATMLFKKGVYLHMP 397

Qy 120 RSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDL 179
Db 398 N-------------------KDGARS--IHTAAAYGHTGIINTLLQKGEKVDVTTNDNY 435

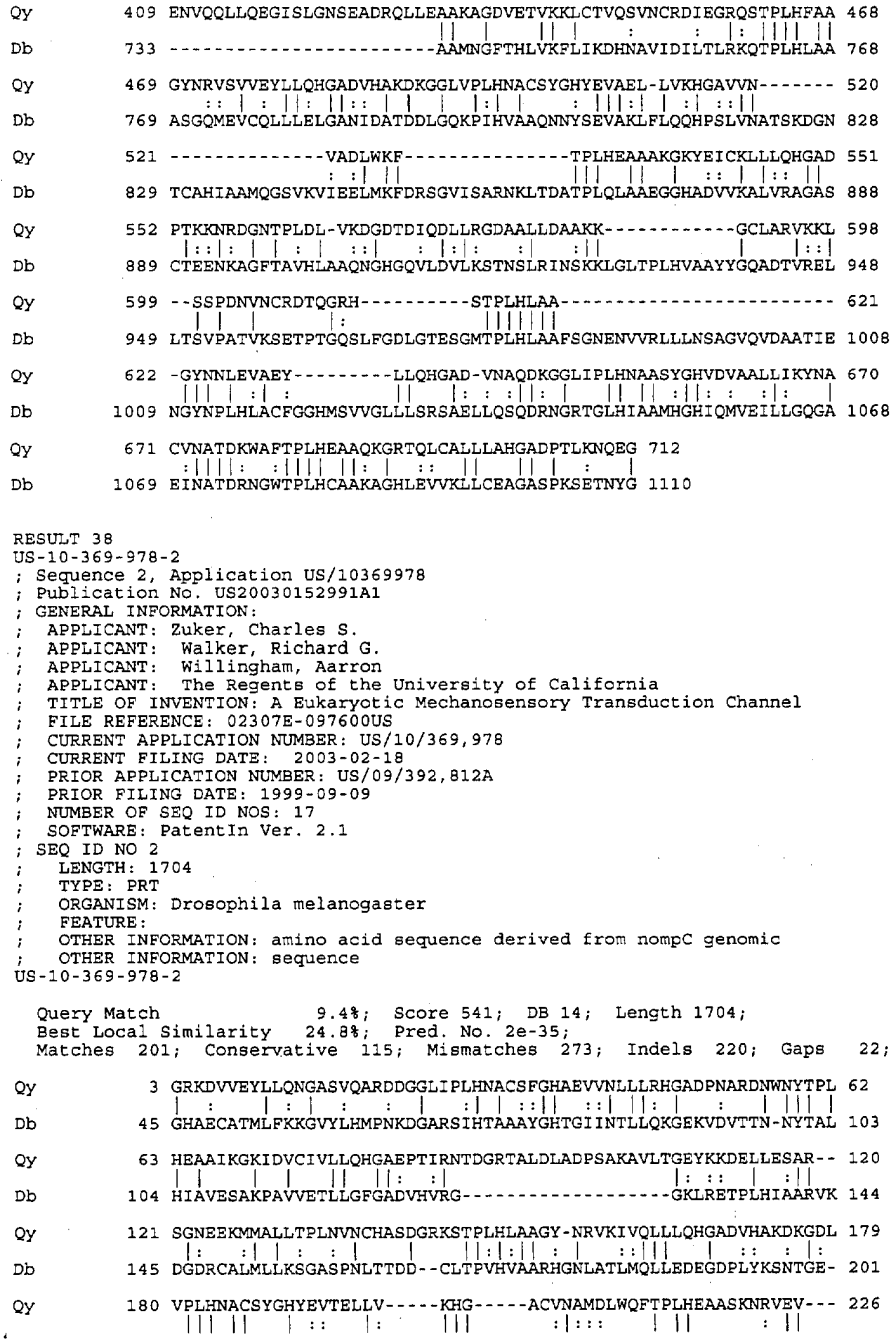
Qy 180 VPLHNACSYGHYEVTELLVKHGACVNAM-DLWQFTPLHEAASKNRVEVCSL-LLSYGADP 237
Db 436 TALHIAVESAKPAVVETLLGFGADVHVRGGKLRETPLHIAARVKDGDRCALMLLKSGASP 495

Qy 238 TLL--NCHNKSAI-----DLAPTPQLKE----RLAYEFKGHSLLQAAREADVTRIKKHLS 286
Db 496 NLTTDDCLTPVHVAARHGNLATLMQLLEDEGDPLYKSNTGETPLHMACRACHPDIVRHL- 554

Qy 287 LEMVNFKH------------PQTHETALH--CAAASPYPK----RKQICELLLRKGANIN 328
Db 555 IETVKEKHGPDKATTYINSVNEDGATALHYTCQITKEEVKIPESDKQIVRMLLENGADVT 614

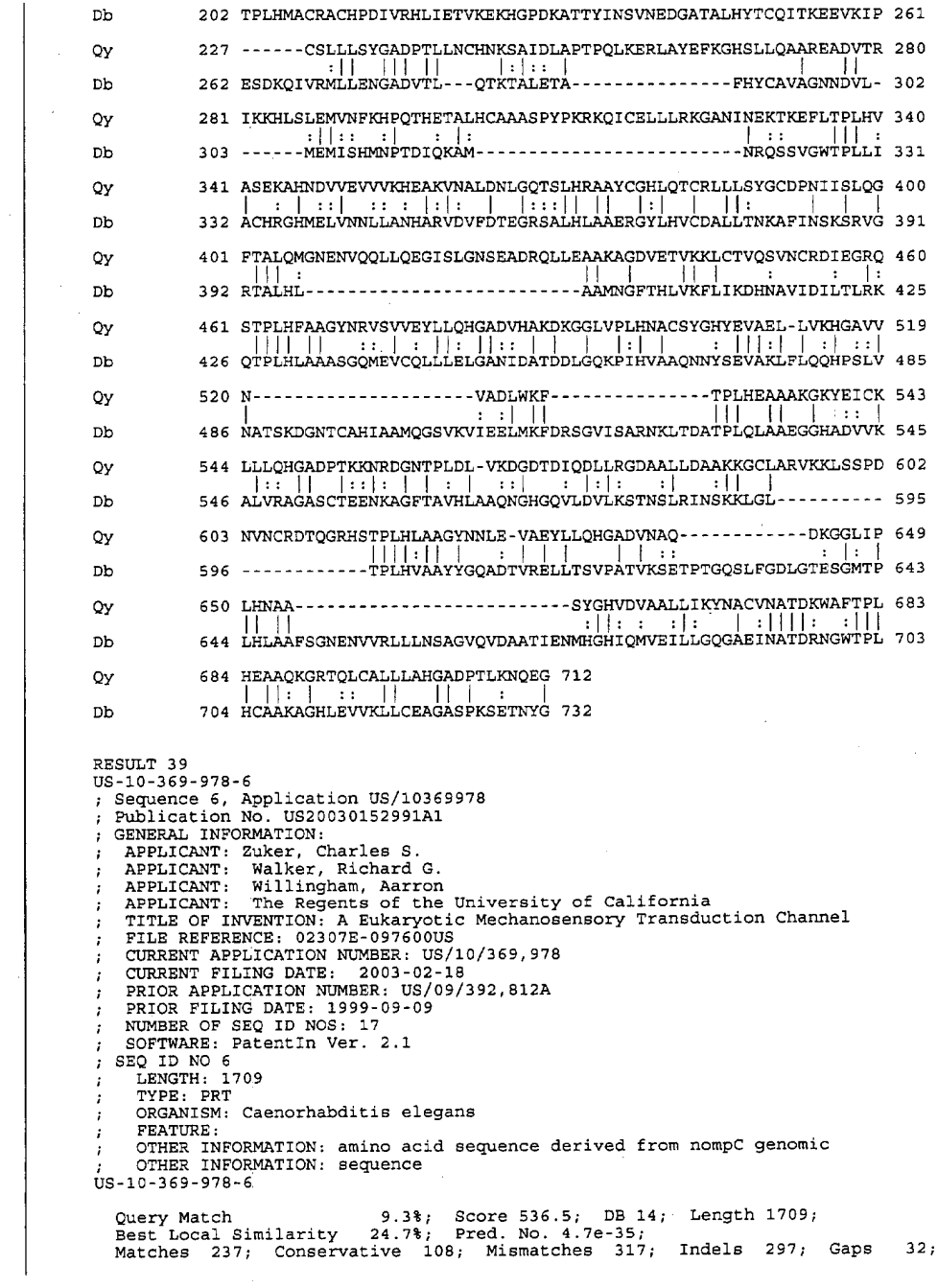
Qy 329 EKTKEFL--TPLHVASEKAHND 348
Db 615 LQTKTALETAFHYCAVAGNNDVLMEMISHMNPTDIQKAMNRQSSVGWTPLLIACHRGHME 674

Qy 349 VVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGN 408
Db 675 LVNNLLANHARVDVFDTEGRSALHLAAERGYLHVCDALLTNKAFINSKSRVGRTALHL-- 732



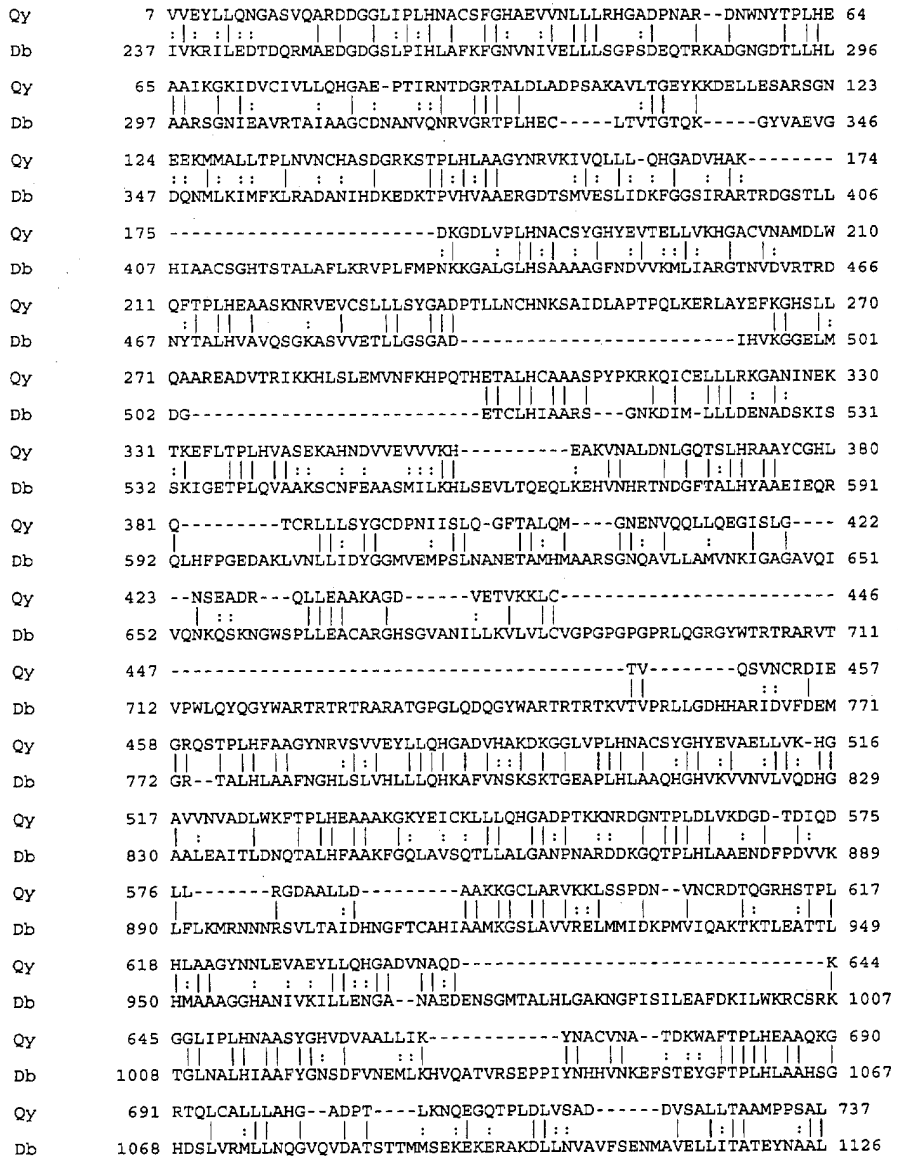
RESULT 38
US-10-369-978-2
; Sequence 2, Application US/10369978
; Publication No. US20030152991A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/10/369,978
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US/09/392,812A
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-10-369-978-2

Query Match 9.4%; Score 541; DB 14; Length 1704;
Best Local Similarity 24.8%; Pred. No. 2e-35;
Matches 201; Conservative 115; Mismatches 273; Indels 220; Gaps 22;



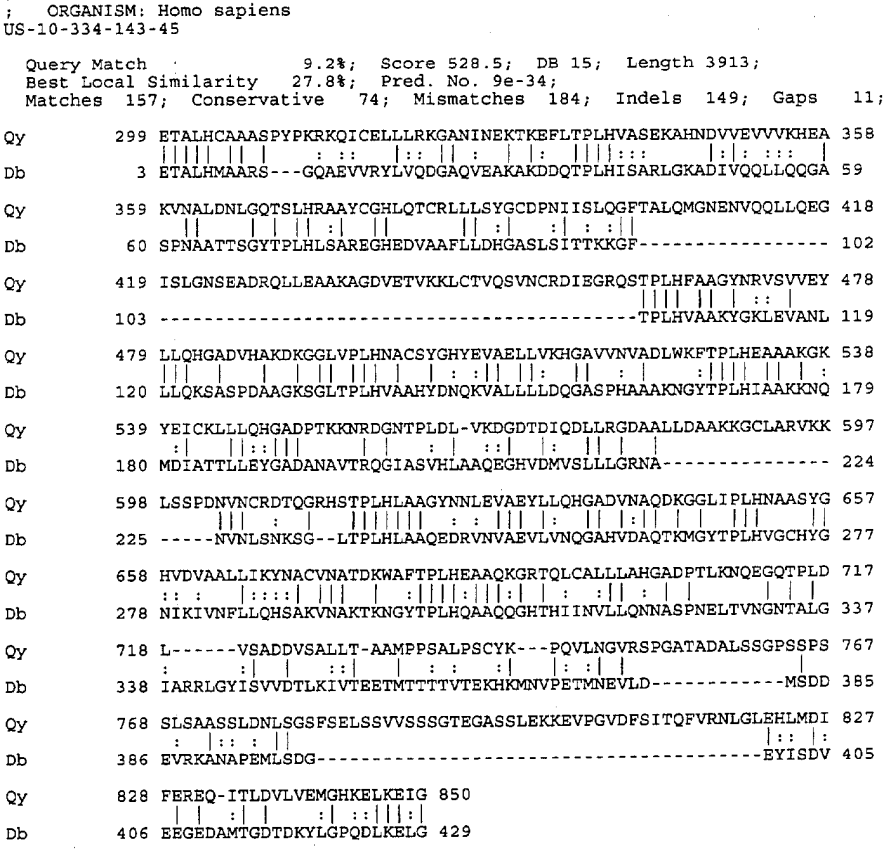
RESULT 39
US-10-369-978-6
; Sequence 6, Application US/10369978
; Publication No. US20030152991A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/10/369,978
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US/09/392,812A
; PRIOR FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1709
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-10-369-978-6

Query Match 9.3%; Score 536.5; DB 14; Length 1709;
Best Local Similarity 24.7%; Pred. No. 4.7e-35;
Matches 237; Conservative 108; Mismatches 317; Indels 297; Gaps 32;



RESULT 40
US-10-334-143-45
; Sequence 45, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-45

Query Match 9.2%; Score 528.5; DB 15; Length 3913;
Best Local Similarity 27.8%; Pred. No. 9e-34;
Matches 157; Conservative 74; Mismatches 184; Indels 149; Gaps 11;



RESULT 41
US-10-094-749-1924
; Sequence 1924, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI

```
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1924
;   LENGTH: 747
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-094-749-1924

  Query Match             8.7%;  Score 501.5;  DB 15;  Length 747;
  Best Local Similarity   23.1%;  Pred. No. 1e-32;
  Matches  210;  Conservative  107;  Mismatches  271;  Indels  321;  Gaps   22;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPN-ARDNWNYTP 61
              |  || : ||| || ::   :||  ||  |   ||   |  |:  ||: |    | ::|
Db         40 GHTDVADVLLQAGADLEHESEGGRTPLMKAARAGHVCTVQFLISKGANVNRTTANNDHTV 99

Qy         62 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 121
              |  |   | : |  :|| |||:|| |  || |                      |:|:|:
Db        100 LSLACAGGHLAVVELLLAHGADPTHRLKDGSTM-------------------LIEAAKG 139

Qy        122 GNEEKMMALL---------------------------------------------------- 131
              |:   :  ||
Db        140 GHTSVVCYLLDYPNNLLSAPPPDVTQLTPPSHDLNRAPRVPVQALPMVVPPQEPDKPPAN 199

Qy        132 --TPLNVNCHASDGRKS----TPLH-LAAGYNRVKI----------VQLLLQHGADVHAK 174
                | | :   |  || |    || | :||  :: :            | :    |: |:
Db        200 VATTLPIRNKAVSGRASAMSNTPTHSIAASISQPQTPTPSPIISPSAMLPIYPAIDIDAQ 259

Qy        175 DKGDL-VPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSY 233
               : :    |  ||: || |: : |:: || :   |  |||| ||:  | |  :||
Db        260 TESNHDTALTLACAGGHEELVQTLLERGASIEHRDKKGFTPLILAATAGHVGVVEILLDN 319

Qy        234 GADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFK 293
              |||                                ::|  |
Db        320 GAD---------------------------------IEAQSE------------------ 328

Qy        294 HPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVV 353
               :| :|  |  | :     |::: |||| :|||   :      ||| :|:   : ::::::
Db        329 --RTKDTPLSLACSG---GRQEVVELLLARGANKEHRNVSDYTPLSLAASGGYVNIIKIL 383

Qy        354 VKHEAKVNAL--DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENV 411
              :  |::|:     || : |  ||  ||    :|||  | | |
Db        384 LNAGAEINSRTGSKLGISPLMLAAMNGHTAAVKLLLDMGSDIN----------------- 426

Qy        412 QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 471
                                                    ||  ::| |  |
Db        427 ----------------------------------------AQIETNRNTALTLACFQG 444

Qy        472 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL--WKFTP 529
              |  ||  ||   |:|  : | || ||  | | |: ||  :|:  || ||   :  : |
Db        445 RTEVVSLLLDRKANVEHRAKTGLTPLMEAASGGYAEVGRVLLDKGADVNAPPVPSSRDTA 504

Qy        530 LHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKK 589
              |  || || |: |:||:  ||    :|: ||
Db        505 LTIAADKGHYKFCELLIGRGAHIDVRNKKGN---------------------------- 535
```

```
Qy        590 GCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIP 649
                                          ||| |||   :|:| : |:| ||||:| |   : |
Db        536 ------------------------TPLWLAANGGHLDVVQLLVQAGADVDAADNRKITP 570

Qy        650 LHNAASYGHVDVAALLIK-------------YNACVNATDKWAFTPLH------------ 684
              |  |   ||| |   |:|              | | :  |||       |
Db        571 LMAAFRKGHVKVVRYLVKEVNQFPSDSECMRYIATI--TDKEMLKKCHLCMESIVQAKDR 628

Qy        685 EAA---------------QKGRTQLCALLLAHGADP-----TLKNQEGQTPLDLVSADDV 724
              :||                :| | :   | ||  :         | :| :  |: : | :
Db        629 QAAEANKNASILLEELDLEKLREESRRLALAAKREKRKEKRRKKKEEQRRKLEEIEAKNK 688

Qy        725 SAL-LTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFS 783
                | ||      |    :|:||    | ||          : :::  :::    |:||
Db        689 ENFELQAAQEKEKLKVEDEPEVL--TEPPSAT---------TTTTIGISATWTTLAGSHG 737

Qy        784 ELSSVVSSS 792
              : :: ::::
Db        738 KRNNTITTT 746
```

```
RESULT 42
US-10-104-047-2529
; Sequence 2529, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20030236392A1el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2529
;   LENGTH: 657
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-104-047-2529

  Query Match             8.2%;  Score 474.5;  DB 15;  Length 657;
  Best Local Similarity   26.7%;  Pred. No. 1.6e-30;
  Matches  169;  Conservative   68;  Mismatches  205;  Indels  191;  Gaps   12;

Qy          4 RKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLH 63
              |:| :  || |||||    |  |   | ||   |  :|||||:  |||    |   :|||
Db        206 REDSIRTLLDNGASVNQCDSNGRTLLANAAYSGSLDVVNLLVSRGADLEIEDAHGHTPLT 265

Qy         64 EAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGN 123
               || :|   |   |:  ||     : || |||  |                    | |:
Db        266 LAARQGHTKVVNCLIGCGANINHTDQDGWTALRSA-------------------AWGGH 305

Qy        124 EEKMMALL-TPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPL 182
               | : |||   : |:|  :| |  | |  ||      ||  ||||||:|:  |      |
Db        306 TEVVSALLYAGVKVDCADADSR--TALRAAAWGGHEDIVLNLLQHGAEVNKADNEGRTAL 363

Qy        183 HNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNC 242
              |   || |: | |: ||| ||  |:   | |  ||      :|
Db        364 IAAAYMGHREIVEHLLDHGAEVNHEDVDGRTALSVAA------LC--------------- 402

Qy        243 HNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETAL 302
                          |              |||:
Db        403 --------VPAS----------KGHA---------------------------------- 410

Qy        303 HCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNA 362
                          :  ||: :|| ::   |: :||| ||: : | |||:::::  | |:
Db        411 -----------SVVSLLIDRGAEVDHCDKDGMTPLLVAAYEGHVDVVDLLLEGGADVDH 458

Qy        363 LDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLG 422
              :| |:|  |  ||  ||       || :|
Db        459 TENNGRTPLLAAASMGHASVVNTLLFWGA---------------------------- 487
```

```
Qy        423 NSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 482
                                          :|:  |:|||  | |  |:    | ||  ||
Db        488 ---------------------------AVDSIDSEGR--TVLSIASAQGNVEVVRTLLDR 518

Qy        483 GADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEIC 542
              | | : ::  |  ||| |   ||  : | |:: ||  | |      |   |: :| |:
Db        519 GLDENHREDAGWTPLHMAAFEGHRLICEALIEQGARTNEIDNDGRIPFILASQEGHYDCV 578

Qy        543 KLLLQHGADPTKKNRDGNTPLDLVK-DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 601
              ::||:: ::  ::  ||   | :   :|  || :||    |
Db        579 QILLENKSNIDQRGYDGRNALRVAALEGHRDIVELLFSHGA------------------ 619

Qy        602 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ 634
               :|||:|  ||:  |:: |  | | :|||  :
Db        620 -DVNCKDADGRPT--LYILALENQLTMAEYFFR 649


RESULT 43
US-10-128-174-3
; Sequence 3, Application US/10128174
; Publication No. US20030199462A1
; GENERAL INFORMATION:
;  APPLICANT: Nunez, Gabriel
;  APPLICANT:  Inohara, Naohiro
;  TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
;  FILE REFERENCE: UM-06967
;  CURRENT APPLICATION NUMBER: US/10/128,174
;  CURRENT FILING DATE:  2002-04-23
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 765
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-128-174-3

  Query Match             8.1%;  Score 465;  DB 14;  Length 765;
  Best Local Similarity   30.1%;  Pred. No. 1.3e-29;
  Matches  144;  Conservative   65;  Mismatches  181;  Indels   88;  Gaps    8;

Qy        245 KSAIDLAP----TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              | |: |:      |: :|   || |     |      :  | :::  |:           ||
Db        338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVEQVRLLLA----------HEV 386

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              : |  || |                              ||| :|::    |:  ::: | |
Db        387 DVDCQTASGY------------------------TPLLIAAQDQQPDLCALLLAHGADA 421

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              | :|  |   || ||  | :| |||| :|   :     :|:| | :  :|
Db        422 NRVDEDGWAPLHFAAQNGDDRTARLLLDHGACVDAQEREGWTPLHLAAQN---------- 471

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              | |   :|| : :|                | |: ||:  |||| || :  ||:|: |
Db        472 --NFENVARLLVSRQA------------DPNLREAEGK--TPLHVAAYFGHVSLVKLLT 514

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||:: |: :    ||| |   |      : |:| ||| :  |   : ||| |||:|||
Db        515 SQGAELDAQQRNLRTPLHLAVERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYL 574

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||:||::||          | ||| |                    || || |  :  | :
Db        575 ICKMLLRYGASLELPTHQGWTPLHL------------------AAYKGHL-EIIHLLA 613

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              : |       : |||||| :    |  ||| ||| || :: |  ||| |
Db        614 ESHANMGALGAVNWTPLHLAARHGEEAVVSALLQCGADPNAAEQSGWTPLHLAVQRSTFL 673

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
                |::::| |:| :|  :|| | || || | :  :|:  ||   :::    ||| |
Db        674 SVINLLEHHANVHARNKVGWTPAHLAALKGNTAILKVLVEAGAQLDVQDGVSCTPLQL 731


RESULT 44
US-10-128-174-34
; Sequence 34, Application US/10128174
; Publication No. US20030199462A1
; GENERAL INFORMATION:
;  APPLICANT: Nunez, Gabriel
;  APPLICANT:  Inohara, Naohiro
;  TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
;  FILE REFERENCE: UM-06967
;  CURRENT APPLICATION NUMBER: US/10/128,174
;  CURRENT FILING DATE:  2002-04-23
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
;   LENGTH: 765
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-128-174-34

  Query Match             8.1%;  Score 465;  DB 14;  Length 765;
  Best Local Similarity   30.1%;  Pred. No. 1.3e-29;
  Matches  144;  Conservative   65;  Mismatches  181;  Indels   88;  Gaps    8;

Qy        245 KSAIDLAP----TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              | |: |:      |: :|   || |     |      :  | :::  |:           ||
Db        338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVEQVRLLLA----------HEV 386

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              : |  || |                              ||| :|::    |:  ::: | |
Db        387 DVDCQTASGY------------------------TPLLIAAQDQQPDLCALLLAHGADA 421

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              | :|  |   || ||  | :| |||| :|   :     :|:| | :  :|
Db        422 NRVDEDGWAPLHFAAQNGDDRTARLLLDHGACVDAQEREGWTPLHLAAQN---------- 471

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              | |   :|| : :|                | |: ||:  |||| || :  ||:|: |
Db        472 --NFENVARLLVSRQA------------DPNLREAEGK--TPLHVAAYFGHVSLVKLLT 514

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||:: |: :    ||| |   |      : |:| ||| :  |   : ||| |||:|||
Db        515 SQGAELDAQQRNLRTPLHLAVERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYL 574

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||:||::||          | ||| |                    || || |  :  | :
Db        575 ICKMLLRYGASLELPTHQGWTPLHL------------------AAYKGHL-EIIHLLA 613

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              : |       : |||||| :    |  ||| ||| || :: |  ||| |
Db        614 ESHANMGALGAVNWTPLHLAARHGEEAVVSALLQCGADPNAAEQSGWTPLHLAVQRSTFL 673

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
                |::::| |:| :|  :|| | || || | :  :|:  ||   :::    ||| |
Db        674 SVINLLEHHANVHARNKVGWTPAHLAALKGNTAILKVLVEAGAQLDVQDGVSCTPLQL 731


RESULT 45
US-10-128-174-35
; Sequence 35, Application US/10128174
; Publication No. US20030199462A1
; GENERAL INFORMATION:
;  APPLICANT: Nunez, Gabriel
;  APPLICANT:  Inohara, Naohiro
;  TITLE OF INVENTION: Methods and Compositions for Regulating Cellular Signaling
;  FILE REFERENCE: UM-06967
;  CURRENT APPLICATION NUMBER: US/10/128,174
;  CURRENT FILING DATE:  2002-04-23
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;   LENGTH: 765
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

US-10-128-174-35

Query Match 8.1%; Score 465; DB 14; Length 765;
Best Local Similarity 30.1%; Pred. No. 1.3e-29;
Matches 144; Conservative 65; Mismatches 181; Indels 88; Gaps 8;

```
Qy        245 KSAIDLAP----TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              | |: |:       |: :|   || |   |      :  | :::  |:           ||
Db        338 KRALQLSDRKNLVPRDEELCIYENKVTPLQFLVAQGSVEQVRLLLA-----------HEV 386

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
               :  |  || |                                |||  :|::    |:  ::: | |
Db        387 DVDCQTASGY------------------------TPLLIAAQDQQPDLCALLLAHGADA 421

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              | :|  |   || ||  |  :| |||| :|   :    :|:| | :  :|
Db        422 NRVDEDGWAPLHFAAQNGDDRTARLLLDHGACVDAQEREGWTPLHLAAQN---------- 471

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              | |   :|| : :|                | |: ||:  |||| || :  ||:|: |
Db        472 --NFENVARLLVSRQA-------------DPNLREAEGK--TPLHVAAYFGHVSLVKLLT 514

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||:: |: :    ||| |   |     : |:| ||| :  |   : ||| |||:|||
Db        515 SQGAELDAQQRNLRTPLHLAVERGKVRAIQHLLKSGAVPDALDQSGYGPLHTAAARGKYL 574

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||:||::||        | ||| |                    || || |  :  | :
Db        575 ICKMLLRYGASLELPTHQGWTPLHL--------------------AAYKGHL-EIIHLLA 613

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              : |       : |||||| :    |   ||| ||| || :: |  ||| |
Db        614 ESHANMGALGAVNWTPLHLAARHGEEAVVSALLQCGADPNAAEQSGWTPLHLAVQRSTFL 673

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
                |::::| |:| :|  :|| | || || | :  :|:  ||   :::      ||| |
Db        674 SVINLLEHHANVHARNKVGWTPAHLAALKGNTAILKVLVEAGAQLDVQDGVSCTPLQL 731
```

Search completed: March 25, 2004, 16:54:32
Job time : 69.188 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:38:10 ; Search time 20.6838 Seconds
(without alignments)
5115.640 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	877.5	15.2	1765	2	T42714	ankyrin 3, splice
2	877.5	15.2	1961	2	T42716	ankyrin 3, splice
3	875.5	15.2	4377	2	A55575	ankyrin 3, long sp
4	873.5	15.1	1943	2	T42713	ankyrin 3, splice
5	871.5	15.1	1940	2	T42715	ankyrin 3, splice
6	860.5	14.9	3924	2	S37431	ankyrin 2, neurona
7	844	14.6	1856	2	B35049	ankyrin 1, erythro
8	844	14.6	1880	2	A35049	ankyrin 1, erythro
9	842	14.6	1881	1	SJHUK	ankyrin 1, erythro
10	839.5	14.6	1848	2	S37771	ankyrin, erythrocy
11	837.5	14.5	1862	2	I49502	ankyrin - mouse
12	834	14.5	1549	2	T13940	ankyrin - fruit fl
13	814.5	14.1	2039	2	T15347	ankyrin-related un
14	628	10.9	791	2	T42691	hypothetical prote
15	545.5	9.5	1411	2	S30355	alpha-latroinsecto
16	514	8.9	934	1	H71274	probable ankyrin -
17	496.5	8.6	1062	2	T30255	inversin - mouse
18	491.5	8.5	1062	2	T14151	Inv protein - mous
19	464.5	8.1	1401	2	S11527	alpha-latrotoxin p
20	456.5	7.9	1188	2	T19552	hypothetical prote
21	451	7.8	1058	2	D82654	ankyrin-like prote
22	431.5	7.5	426	2	AE2149	hypothetical prote
23	426	7.4	991	2	T25412	hypothetical prote
24	405.5	7.0	1184	2	T00253	gene Ankhzn protei
25	388	6.7	397	2	T46445	hypothetical prote
26	372.5	6.5	2584	2	T24158	hypothetical prote
27	372.5	6.5	2606	2	T24157	hypothetical prote
28	363.5	6.3	1031	2	T43458	hypothetical prote
29	350	6.1	557	2	T46507	hypothetical prote
30	348.5	6.0	606	2	AC2508	hypothetical prote
31	345	6.0	1423	1	I37275	death-associated p
32	336.5	5.8	1107	2	T21280	hypothetical prote
33	321	5.6	1398	2	T21884	hypothetical prote
34	310	5.4	1435	2	T32930	hypothetical prote
35	306	5.3	1322	2	A59288	myosin heavy chain
36	305	5.3	679	2	B45771	2-5A-dependent RNA
37	300.5	5.2	1016	2	T19006	ankyrin related pr
38	299.5	5.2	633	2	T27499	hypothetical prote
39	299	5.2	247	2	D84448	probable ankyrin [
40	294.5	5.1	900	2	A42024	transcription fact
41	290	5.0	1004	2	A55142	myosin-light-chain
42	287.5	5.0	656	2	A34793	sex-determining pr
43	287	5.0	1295	2	T21720	hypothetical prote
44	285	4.9	1083	2	S48460	probable membrane
45	284	4.9	237	2	T50984	related to 26s pro

ALIGNMENTS

RESULT 1
T42714
ankyrin 3, splice form 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42714
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene f
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1765 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710551; PIDN:AAB01605.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 1587/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 15.2%; Score 877.5; DB 2; Length 1765;
Best Local Similarity 30.9%; Pred. No. 1.9e-46;
Matches 277; Conservative 125; Mismatches 377; Indels 117; Gaps 22;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:| |     |  || |  | |||| :|: :||: ||:    :|||
Db         68 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 127

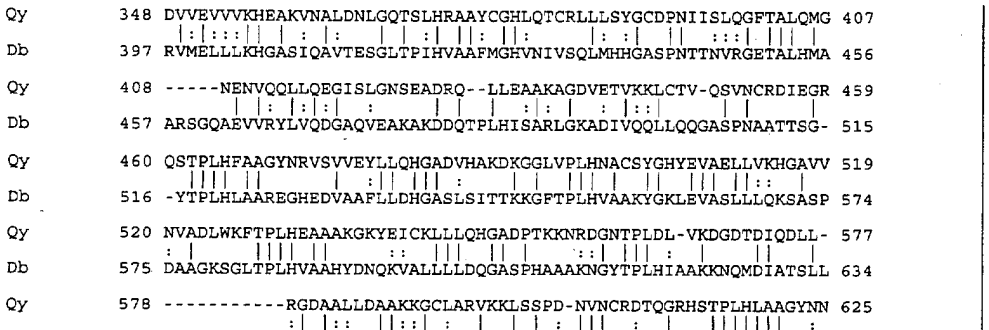
Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :||  ::   || | | :|     |   :|  : |   :
Db        128 YMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 187

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |  |   :: |  |||   : |          ||||:|| |  : :  |||   | |
Db        188 ALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 247

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |   |:  : :||:  || ::|       ||||  |      :|  :||
Db        248 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRS 307

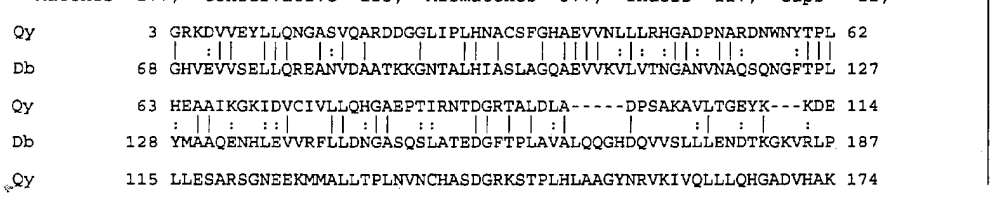
Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |    |                 | | |  | : |       ||:   :  :|
Db        308 A-PILSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 339

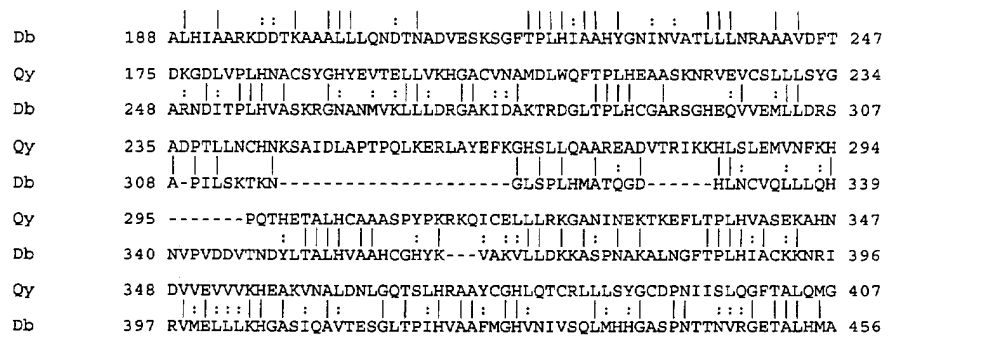
Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | |: | |     ||||:| :|
Db        340 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKASPNAKALNGFTPLHIACKKNRI 396
```



RESULT 2
T42716
ankyrin 3, splice form 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42716
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42716
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1961 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 15.2%; Score 877.5; DB 2; Length 1961;
Best Local Similarity 30.9%; Pred. No. 2.2e-46;
Matches 277; Conservative 125; Mismatches 377; Indels 117; Gaps 22;





RESULT 3
A55575
ankyrin 3, long splice form - human
N;Alternate names: ankyrin G
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 20-Sep-1999
C;Accession: A55575
R;Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A;Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the axo
A;Reference number: A55575; MUID:95138209; PMID:7836469
A;Accession: A55575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4377 <KOR>
A;Cross-references: GB:U13616; NID:g608024; PIDN:AAA64834.1; PID:g608025
C;Genetics:
A;Gene: GDB:ANK3
A;Cross-references: GDB:424503; OMIM:600465
A;Map position: 10q21-10q21
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

```
C;Keywords: alternative splicing; peripheral membrane protein
F;73-105/Domain: ankyrin repeat homology <AN01>
F;106-138/Domain: ankyrin repeat homology <AN02>
F;139-171/Domain: ankyrin repeat homology <AN03>
F;172-200/Domain: ankyrin repeat homology <AN04>
F;201-233/Domain: ankyrin repeat homology <AN05>
F;234-266/Domain: ankyrin repeat homology <AN06>
F;267-299/Domain: ankyrin repeat homology <AN07>
F;300-332/Domain: ankyrin repeat homology <AN08>
F;333-365/Domain: ankyrin repeat homology <AN09>
F;366-398/Domain: ankyrin repeat homology <AN10>
F;399-431/Domain: ankyrin repeat homology <AN11>
F;432-464/Domain: ankyrin repeat homology <AN12>
F;465-497/Domain: ankyrin repeat homology <AN13>
F;498-530/Domain: ankyrin repeat homology <AN14>
F;531-563/Domain: ankyrin repeat homology <AN15>
F;564-596/Domain: ankyrin repeat homology <AN16>
F;597-629/Domain: ankyrin repeat homology <AN17>
F;630-662/Domain: ankyrin repeat homology <AN18>
F;663-695/Domain: ankyrin repeat homology <AN19>
F;696-728/Domain: ankyrin repeat homology <AN20>
F;729-761/Domain: ankyrin repeat homology <AN21>
F;762-794/Domain: ankyrin repeat homology <AN22>
F;795-827/Domain: ankyrin repeat homology <AN23>

  Query Match             15.2%;  Score 875.5;  DB 2;  Length 4377;
  Best Local Similarity   30.5%;  Pred. No. 1.1e-45;
  Matches  273;  Conservative  126;  Mismatches  362;  Indels  135;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:| |    |  || |  | |||| :|: :||: ||:    :|||
Db         85 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 144

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :||  ::   || | | :|     |     :|  : |   :
Db        145 YMAAQENHLEVVKFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 204

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |  |   ::  |  |||   : |          ||||:|| |  : :  |||    | |
Db        205 ALHIAARKDDTKAAALLLQNDNNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 264

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |   |:  : :||:  || ::|      ||||  |      :|  :||
Db        265 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRA 324

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |   |                 | | |  : |     ||:   :  :|
Db        325 A-PILSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 356

Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | || | |     ||||:| :|
Db        357 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKANPNAKALNGFTPLHIACKKNRI 413

Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
              |:|:::|| | : |:   | | :| ||: ||:     |: :|  ||  :::| |||| |
Db        414 KVMELLLKHGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNTTNVRGETALHMA 473

Qy        408 -----NENVQQLLQEGISLGNSEADRQ--LLEAAKAGDVETVKKLCTV-QSVNCRDIEGR 459
                  | |: |:|:|  :     | |  |  :|: |  : |::|     | |     |
Db        474 ARSGQAEVVRYLVQDGAQVEAKAKDDQTPLHISARLGKADIVQQLLQQGASPNAATTSG- 532

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
              |||| :|     |  :|| ||| :    | |  ||| |  ||  ||| ||::  |
Db        533 -YTPLHLSAREGHEDVAAFLLDHGASLSITTKKGFTPLHVAAKYGKLEVANLLLQKSASP 591

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL- 577
              : |    |||| ||     ::  |||  || |    ::| ||| :  |     ||
Db        592 DAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAAKNGYTPLHIAAKKNQMDIATTLL 651

Qy        578 ----------RGDAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNN 625
                       :| |::  ||::| :   |    : |||  :  |   |||||||  :
Db        652 EYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNANVNLSNKSG--LTPLHLAAQEDR 709

Qy        626 LEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHE 685
              : ||| |:  || |:|| | |  |||    ||:: :   |::::| |||  |  :||||:
Db        710 VNVAEVLVNQGAHVDAQTKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTPLHQ 769

Qy        686 AAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL------VSADDVSALLT-AAMPPSALP 738
              |||:| | :  :|| : | |       | | | :      :|  |    ::|    |  : :
Db        770 AAQQGHTHIINVLLQNNASPNELTVNGNTALGIARRLGYISVVDTLKIVTEETMTTTTVT 829

Qy        739 SCYK---PQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 795
              :|    |: :| |                |   :  |:: : ||
Db        830 EKHKMNVPETMNEVLD------------MSDDEVRKANAPEMLSDG-------------- 863

Qy        796 GASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQ-ITLDVLVEMGHKELKEIG 850
                                          |:: |: | |  :| |    :| ::|||:|
Db        864 --------------------------EYISDVEEGEDAMTGDTDKYLGPQDLKELG 893


RESULT 4
T42713
ankyrin 3, splice form 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42713
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L.
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene f
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42713
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1943 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 855/1
C;Function:
A;Description: supposed to play an important role in the polarized distribution of many i
A;Note: major kidney ankyrin
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

  Query Match             15.1%;  Score 873.5;  DB 2;  Length 1943;
  Best Local Similarity   30.6%;  Pred. No. 3.9e-46;
  Matches  274;  Conservative  122;  Mismatches  365;  Indels  135;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |||  |:| |    |  || |  | |||| :|: :||: ||:    :|||
Db         68 GHVEVVSELLQREANVDAATKKGNTALHIASLAGQAEVVKVLVTNGANVNAQSQNGFTPL 127

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114
              : || :  ::|   || :||  ::   || | | :|     |     :|  : |   :
Db        128 YMAAQENHLEVVRFLLDNGASQSLATEDGFTPLAVALQQGHDQVVSLLLENDTKGKVRLP 187

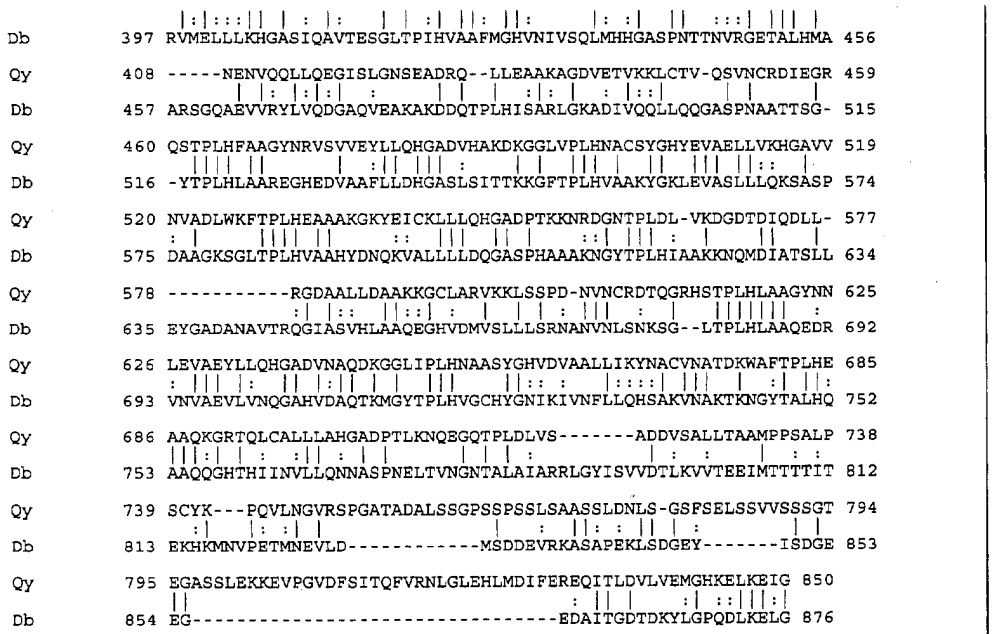
Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
              |  |   ::  |  |||   : |          ||||:|| |  : :  |||    | |
Db        188 ALHIAARKDDTKAAALLLQNDTNADVESKSGFTPLHIAAHYGNINVATLLLNRAAAVDFT 247

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
              : |: ||| |   |:  : :||:  || ::|      ||||  |      :|  :||
Db        248 ARNDITPLHVASKRGNANMVKLLLDRGAKIDAKTRDGLTPLHCGARSGHEQVVEMLLDRS 307

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              | | |   |                 | | |  : |     ||:   :  :|
Db        308 A-PILSKTKN--------------------GLSPLHMATQGD------HLNCVQLLLQH 339

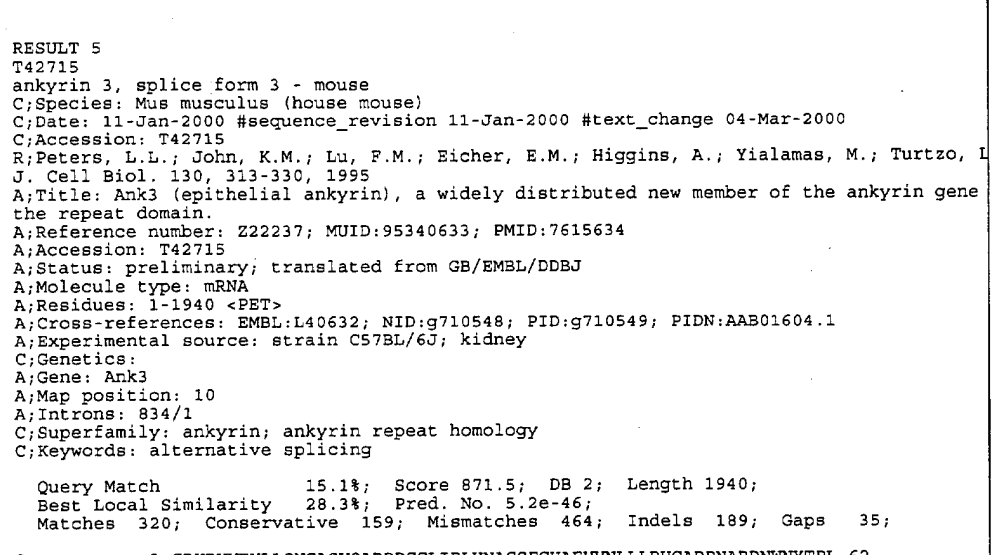
Qy        295 -------PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
                    : |||| ||   : |   : ::|| | |: | |     ||||:| :|
Db        340 NVPVDDVTNDYLTALHVAAHCGHYK---VAKVLLDKKASPNAKALNGFTPLHIACKKNRI 396

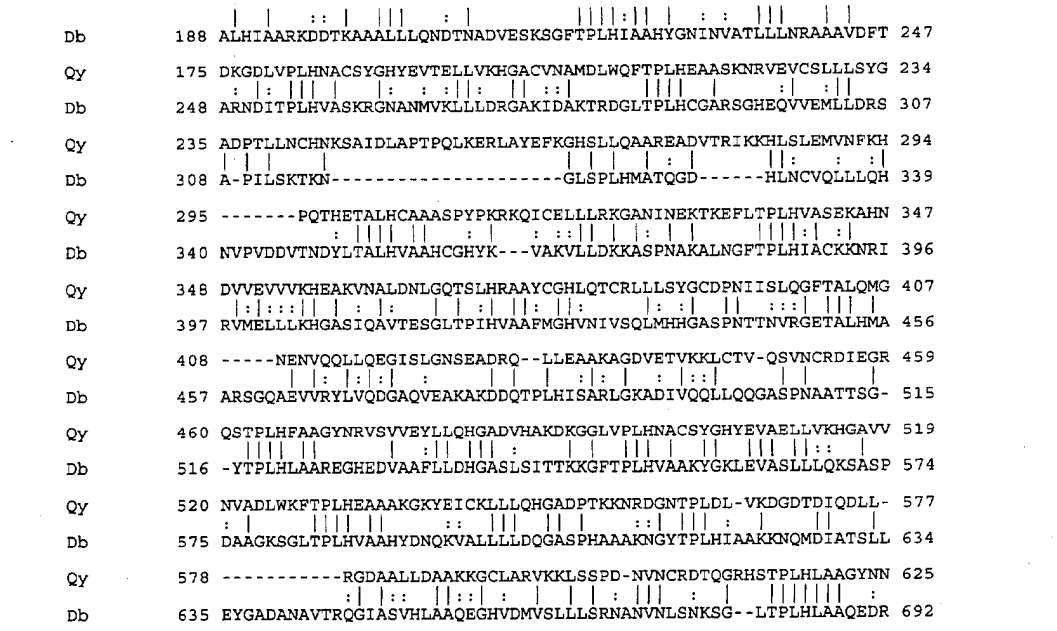
Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMG 407
```



RESULT 5
T42715
ankyrin 3, splice form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T42715
R;Peters, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, L
J. Cell Biol. 130, 313-330, 1995
A;Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene
the repeat domain.
A;Reference number: Z22237; MUID:95340633; PMID:7615634
A;Accession: T42715
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1940 <PET>
A;Cross-references: EMBL:L40632; NID:g710548; PID:g710549; PIDN:AAB01604.1
A;Experimental source: strain C57BL/6J; kidney
C;Genetics:
A;Gene: Ank3
A;Map position: 10
A;Introns: 834/1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing

Query Match 15.1%; Score 871.5; DB 2; Length 1940;
Best Local Similarity 28.3%; Pred. No. 5.2e-46;
Matches 320; Conservative 159; Mismatches 464; Indels 189; Gaps 35;

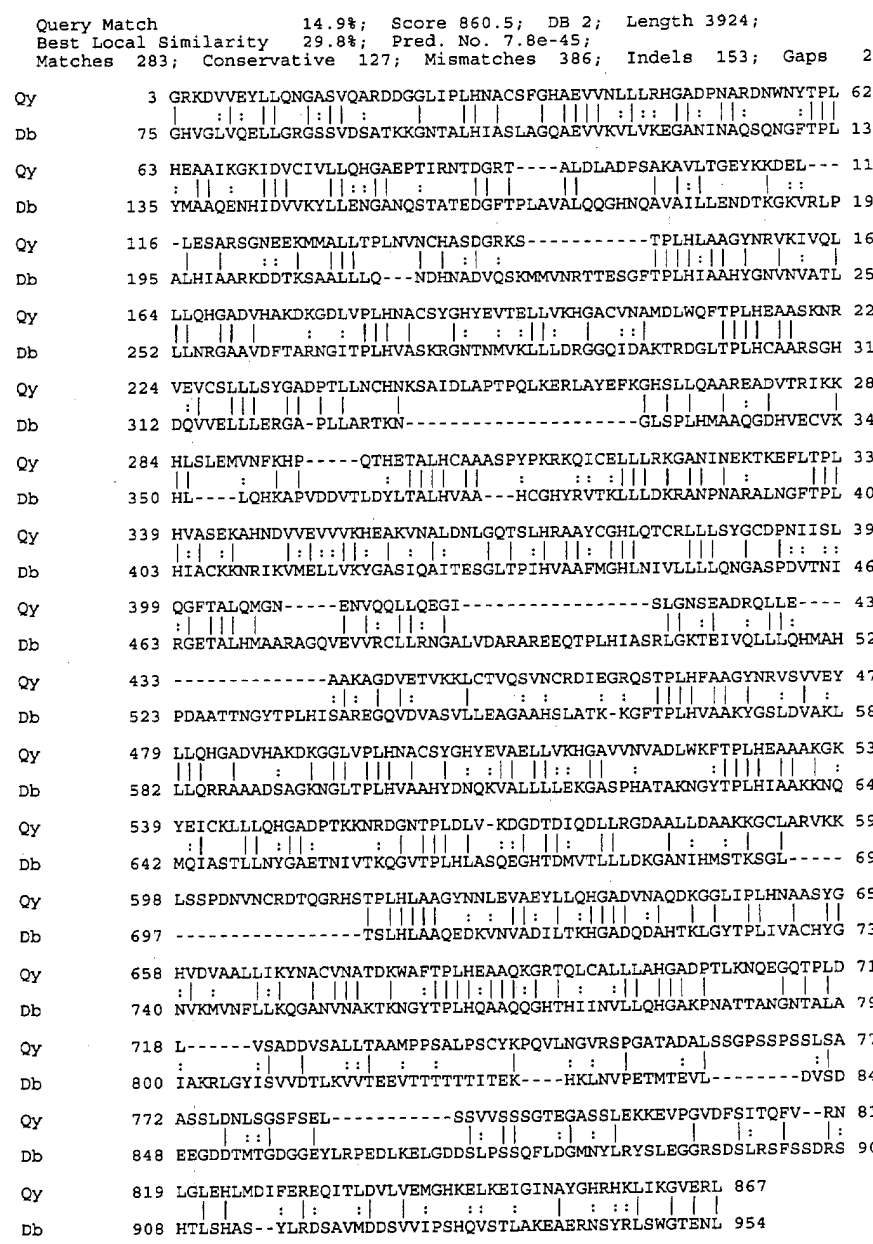




RESULT 6
S37431
ankyrin 2, neuronal long splice form - human
N;Alternate names: ankyrin B, 440K splice form; ankyrin-B; brain ankyrin; non-erythroid a
N;Contains: ankyrin 2, short form

C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C;Accession: S37431; A39643; B39643; A40334; A49462; S14533; S14569
R;Chan, W.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37431
A;Accession: S37431
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-3924 <CHA>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
R;Otto, E.; Kunimoto, M.; McLaughlin, T.; Bennett, V.
J. Cell Biol. 114, 241-253, 1991
A;Title: Isolation and characterization of cDNAs encoding human brain ankyrins reveal a
A;Reference number: A39643; MUID:91302466; PMID:1830053
A;Accession: A39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2077 <OT1>
A;Cross-references: GB:X56957
A;Accession: B39643
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1443,3585-3924 <OTT>
A;Cross-references: EMBL:X56958
R;Tse, W.T.; Menninger, J.C.; Yang-Feng, T.L.; Francke, U.; Sahr, K.E.; Lux, S.E.; Ward,
Genomics 10, 858-866, 1991
A;Title: Isolation and chromosomal localization of a novel nonerythroid ankyrin gene.
A;Reference number: A40334; MUID:92009921; PMID:1833308
A;Accession: A40334
A;Molecule type: DNA
A;Residues: 463-474,'PE',477-495 <TSE>
A;Cross-references: GB:M37123; NID:g178647; PIDN:AAA62828.1; PID:g178648
R;Chan, W.; Kordeli, E.; Bennett, V.
J. Cell Biol. 123, 1463-1473, 1993
A;Title: 440-kD ankyrinB: structure of the major developmentally regulated domain and se
A;Reference number: A49462; MUID:94075409; PMID:8253844
A;Accession: A49462
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3924 <RES>
A;Cross-references: EMBL:Z26634; NID:g406287; PIDN:CAA81387.1; PID:g406288
C;Genetics:
A;Gene: GDB:ANK2
A;Cross-references: GDB:127607; OMIM:106410
A;Map position: 4q25-4q27
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-3924/Product: ankyrin 2, long form #status predicted <MAT>
F;2-1443,3585-3924/Product: ankyrin 2, short form #status predicted <MA2>
F;63-95/Domain: ankyrin repeat homology <AN01>
F;96-128/Domain: ankyrin repeat homology <AN02>
F;129-161/Domain: ankyrin repeat homology <AN03>
F;162-190/Domain: ankyrin repeat homology <AN04>
F;191-223/Domain: ankyrin repeat homology <AN05>
F;232-264/Domain: ankyrin repeat homology <AN06>
F;265-297/Domain: ankyrin repeat homology <AN07>
F;298-330/Domain: ankyrin repeat homology <AN08>
F;331-363/Domain: ankyrin repeat homology <AN09>
F;364-396/Domain: ankyrin repeat homology <AN10>
F;397-429/Domain: ankyrin repeat homology <AN11>
F;430-462/Domain: ankyrin repeat homology <AN12>
F;463-495/Domain: ankyrin repeat homology <AN13>
F;496-528/Domain: ankyrin repeat homology <AN14>
F;529-561/Domain: ankyrin repeat homology <AN15>
F;562-594/Domain: ankyrin repeat homology <AN16>
F;595-627/Domain: ankyrin repeat homology <AN17>
F;628-660/Domain: ankyrin repeat homology <AN18>
F;661-693/Domain: ankyrin repeat homology <AN19>
F;694-726/Domain: ankyrin repeat homology <AN20>
F;727-759/Domain: ankyrin repeat homology <AN21>
F;760-792/Domain: ankyrin repeat homology <AN22>
F;793-825/Domain: ankyrin repeat homology <AN23>

Query Match 14.9%; Score 860.5; DB 2; Length 3924;
Best Local Similarity 29.8%; Pred. No. 7.8e-45;
Matches 283; Conservative 127; Mismatches 386; Indels 153; Gaps 21;



RESULT 7
B35049
ankyrin 1, erythrocyte splice form 3 - human

```
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C;Accession: B35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:90175370; PMID:1689849
A;Accession: B35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1856 <LAM>
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

  Query Match             14.6%;  Score 844;  DB 2;  Length 1856;
  Best Local Similarity   25.7%;  Pred. No. 2.6e-44;
  Matches  281;  Conservative  151;  Mismatches  355;  Indels  306;  Gaps   28;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |      :|||
Db         89 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 148

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                |  :|  :|   |: :| :  :          || | |||          |||
Db        149 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 199

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASD--------GRKS----------TPLHLAAGY 155
              ::|  :::|         |||::  |  :          | |          ||||:|:
Db        200 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGSSVNFTPQNGITPLHIASRR 249

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
                | :|:|||  || :  | | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        250 GNVIMVRLLLDRGAQIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:          ||
Db        310 HMAAQGDHLDCVRLLLQYDAE-----------ID------------------------- 332

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                | | || ||   :    :: ::|| |||  | :
Db        333 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 372

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : || |  ||:
Db        373 TPLHIACKKNHVRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 432

Qy        396 ISLQGFTALQMG-----NENVQQLLQE----------------GISLGNSEADRQLLE- 432
              :::  | | |              | : |||                :|::   : |||
Db        433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLEN 492

Qy        433 -----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
                              ||: | ||||  |   : |  |   :|   |||| || | :|
Db        493 NANPNLATTAGHTPLHIAARFGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVR 550

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
              | | ||:  |  :|  | || ||| |  : : :: :||:  |   : :  |  :|||| |
Db        551 VAELLLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 609

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
              | : : |: : |||:|      ::  | ||| |  ::|  ::  ||
Db        610 AKQNQVEVARSLLQYGGSANAESVQGVTPLHLAAQEGHAEMVALL---------------- 654

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
                    ||   | |  :  |  |||||  |    :: ||: |::||  |:|  : |  |||
Db        655 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHV 707

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
              |: ||:: :   |::: | |||  |  ::|||:|||:| | :  ||| :|| |   : :|
Db        708 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDG 767

Qy        713 QTPLDL--------------VSADDVSALLTAAMPPSALPSC----------------- 740
                |||  :            |  |: | :| :     : |
Db        768 TTPLAIAKRLGYISVTDVLKVVTDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISF 827

Qy        741 -----------------YKPQVLNGVRSPGATADALSSGPSSPSSLS-----AASSLDN 777
                              : |::   | ||           |  | ::       :   :
Db        828 KAERRDSRDVDEEKELLDFVPKLDQVVESPAI--------PRIPCAMPETVVIRSEEQEQ 879

Qy        778 LSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIPEREQITLDV 837
              |  : | | : ||  || : ::      |      :  |:
Db        880 ASKEYDEDSLIPSSPATETSDNISPVASPVHTGFLVSFM---------------------- 918

Qy        838 LVEMGHKELKEIGINAYGHRHKLIKGVE-------------RLISGQQGLNPYLTLNTSG 884
                        :   | :  | ||  :: |              ||:  |:   |      |
Db        919 --------VDARGGSMRGSRHNGLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEG 970

Qy        885 SGTILIDLSP--------------------DDKEFQSVEEEMQSTVREHRDGGHAGGIF 923
                : :| | |                    |:|   :  |  | :|||
Db        971 LASRIIALGPTGAQFLSPVIVEIPHFASHGRGDRELVVLRSENGSVWKEHR--------- 1021

Qy        924 NRYNILKIQKVCN 936
              :||    : :: |
Db       1022 SRYGESYLDQILN 1034


RESULT 8
A35049
ankyrin 1, erythrocyte splice form 2 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C;Accession: A35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; Ka
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:90175370; PMID:1689849
A;Accession: A35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1880 <LAM>
A;Cross-references: GB:M28880
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
```

```
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; cytoskeleton
F;2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F;2-1513,1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>

  Query Match             14.6%;  Score 844;  DB 2;  Length 1880;
  Best Local Similarity   25.7%;  Pred. No. 2.6e-44;
  Matches  281;  Conservative  151;  Mismatches  355;  Indels  306;  Gaps   28;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         89 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 148

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
Db        149 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 199

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASD--------GRKS----------TPLHLAAGY 155
Db        200 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGSSVNFTPQNGITPLHIASRR 249

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
Db        250 GNVIMVRLLLDRGAQIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
Db        310 HMAAQGDHLDCVRLLLQYDAE-----------ID-------------------------- 332

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
Db        333 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 372

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
Db        373 TPLHIACKKNHVRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 432

Qy        396 ISLQGFTALQMG-----NENVQQLLQE----------------GISLGNSEADRQLLE- 432
Db        433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLEN 492

Qy        433 ----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
Db        493 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVR 550

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
Db        551 VAELLLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 609

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
Db        610 AKQNQVEVARSLLQYGGSANAESVQGVTPLHLAAQEGHAEMVALL---------------- 654

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
Db        655 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHV 707

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
Db        708 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDG 767

Qy        713 QTPLDL--------------VSADDVSALLTAAMPPSALPSC------------------ 740
Db        768 TTPLAIAKRLGYISVTDVLKVVTDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISF 827

Qy        741 ------------------YKPQVLNGVRSPGATADALSSGPSSPSSLS-----AASSLDN 777
Db        828 KAERRDSRDVDEEKELLDFVPKLDQVVESPAI--------PRIPCAMPETVVIRSEEQEQ 879

Qy        778 LSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDV 837
Db        880 ASKEYDEDSLIPSSPATETSDNISPVASPVHTGFLVSFM---------------------- 918

Qy        838 LVEMGHKELKEIGINAYGHRHKLIKGVE------------RLISGQQGLNPYLTLNTSG 884
Db        919 --------VDARGGSMRGSRHNGLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEG 970

Qy        885 SGTILIDLSP--------------------DDKEFQSVEEEMQSTVREHRDGGHAGGIF 923
Db        971 LASRIIALGPTGAQFLSPVIVEIPHFASHGRGDRELVVLRSENGSVWKEHR--------- 1021

Qy        924 NRYNILKIQKVCN 936
Db       1022 SRYGESYLDQILN 1034


RESULT 9
SJHUK
ankyrin 1, erythrocyte splice form 1 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999
C;Accession: S08275; A33219; PC2220; A35443
R;Lux, S.E.; John, K.M.; Bennett, V.
Nature 344, 36-42, 1990
A;Title: Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure wi
A;Reference number: S08275; MUID:90158830; PMID:2137557
A;Accession: S08275
A;Molecule type: mRNA
A;Residues: 1-1881 <LU1>
A;Cross-references: EMBL:X16609; NID:g28701; PIDN:CAA34610.1; PID:g28702
A;Accession: A33219
A;Molecule type: protein
A;Residues: 2-7,'X',9-17,'X',19-20,'T',22-30;733-749,'A',751-753;828-833,'X',835-855,'X',
X',1367;1383-1427;1601-1630;1686-1698,'D',1700;1763-1772 <LUX>
A;Note: 845-Arg and 1392-Thr were also found
R;Hermann, J.; Barel, M.; Frade, R.
Biochem. Biophys. Res. Commun. 204, 453-460, 1994
A;Title: Human erythrocyte ankyrin, a cytoskeleton component, generates the p57 membrane
A;Reference number: PC2220; MUID:95071348; PMID:7526850
A;Accession: PC2220
A;Molecule type: protein
A;Residues: 910-929 <HER>
R;Davis, L.H.; Bennett, V.
J. Biol. Chem. 265, 10589-10596, 1990
A;Title: Mapping the binding sites of human erythrocyte ankyrin for the anion exchanger a
A;Reference number: A35443; MUID:90285190; PMID:2141335
A;Accession: A35443
A;Molecule type: protein
A;Residues: 'X',5,'X',7-12;403-417,'X',419-422,'H',424,'LQ';797-800,'L',802-814;862-863,'
C;Genetics:
```

```
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 8p11.2-8p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing; phosphoprotein
F;2-1881/Product: ankyrin 1, erythrocyte form 1 #status predicted <MAT1>
F;2-1512,1675-1881/Product: ankyrin 2.2, erythrocyte #status predicted <MAT2>
F;2-827/Domain: 89K #status predicted <DOM1>
F;2-827/Region: anion exchange protein binding
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-109/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>
F;436-468/Domain: ankyrin repeat homology <AN13>
F;469-501/Domain: ankyrin repeat homology <AN14>
F;502-534/Domain: ankyrin repeat homology <AN15>
F;535-567/Domain: ankyrin repeat homology <AN16>
F;568-600/Domain: ankyrin repeat homology <AN17>
F;601-633/Domain: ankyrin repeat homology <AN18>
F;634-666/Domain: ankyrin repeat homology <AN19>
F;667-699/Domain: ankyrin repeat homology <AN20>
F;700-732/Domain: ankyrin repeat homology <AN21>
F;733-765/Domain: ankyrin repeat homology <AN22>
F;766-798/Domain: ankyrin repeat homology <AN23>
F;828-1382/Domain: 62K #status predicted <DOM2>
F;828-1382/Region: spectrin binding
F;1383-1881/Domain: 55K #status predicted <DOM3>

  Query Match           14.6%;  Score 842;  DB 1;  Length 1881;
  Best Local Similarity  25.5%;  Pred. No. 3.5e-44;
  Matches  279;  Conservative  152;  Mismatches  356;  Indels  306;  Gaps   27;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |     :|||
Db         89 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 148

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                |  :|  :|   |: :| :  :        || | |||         |||
Db        149 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 199

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
              ::|  :::|         |||::  |  :   :                 ||||:|:
Db        200 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 249

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
                | :|:|||  || :  | | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        250 GNVIMVRLLLDRGAQIETKTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 309

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:          ||
Db        310 HMAAQGDHLDCVRLLLQYDAE----------ID------------------------- 332

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                | | || ||   :    :: ::|| |||   | :
Db        333 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 372

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : ||  |  ||:
Db        373 TPLHIACKKNHVRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 432

Qy        396 ISLQGFTALQMG-----NENVQQLLQE-----------------GISLGNSEADRQLLE- 432
              ::  | | |       | : |||                   :|::   : |||
Db        433 SNVKVETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQTPLHCAARIGHTNMVKLLLEN 492

Qy        433 -----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
                                ||: | ||||  |   : |  |   :|   |||| || | :|
Db        493 NANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKG--FTPLHVAAKYGKVR 550

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
              | | ||:  |  :|  | || ||| |  : : :: :||:  |    : :  |  :|||| |
Db        551 VAELLLERDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 609

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
              | : : |: : |||:|      ::  | ||| |  ::|  ::  ||
Db        610 AKQNQVEVARSLLQYGGSANAESVQGVTPLHLAAQEGHAEMVALL--------------- 654

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
                ||   | |  :  |  |||||  |   :: ||: |::||  |:|  : |  |||
Db        655 -----LSKQANGNLGNKSG--LTPLHLVAQEGHVPVADVLIKHGVMVDATTRMGYTPLHV 707

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
              |: ||:: :   |::: | |||  |  ::|||:|||:| | :  ||| :|| |   : :|
Db        708 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSDG 767

Qy        713 QTPLDL--------------VSADDVSALLTAAMPPSALPSC------------------ 740
              |||  :                |  |: | :|  :     :  |
Db        768 TTPLAIAKRLGYISVTDVLKVVTDETSFVLVSDKHRMSFPETVDEILDVSEDEGEELISF 827

Qy        741 ------------------YKPQVLNGVRSPGATADALSSGPSSPSSLS-----AASSLDN 777
                               : |::  | ||          |  | ::     :   :
Db        828 KAERRDSRDVDEEKELLDFVPKLDQVVESPAI--------PRIPCAMPETVVIRSEEQEQ 879

Qy        778 LSGGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDV 837
              |  : | | : ||  || : ::      |     :  |:
Db        880 ASKEYDEDSLIPSSPATETSDNISPVASPVHTGFLVSFM--------------------- 918

Qy        838 LVEMGHKELKEIGINAYGHRHKLIKGVE-------------RLISGQQGLNPYLTLNTSG 884
                    :  | :  | ||  :: |                 ||:  |:   |      |
Db        919 --------VDARGGSMRGSRHNGLRVVIPPRTCAAPTRITCRLVKPQKLSTPPPLAEEEG 970

Qy        885 SGTILIDLSP--------------------DDKEFQSVEEEMQSTVREHRDGGHAGGIF 923
                : :| | |                    |:|   :  |  |  :|||
Db        971 LASRIIALGPTGAQFLSPVIVEIPHFASHGRGDRELVVLRSENGSVWKEHR--------- 1021

Qy        924 NRYNILKIQKVCN 936
              :||    : :: |
Db       1022 SRYGESYLDQILN 1034

RESULT 10
S37771
ankyrin, erythrocyte - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: S37771
R;Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found am
A;Reference number: S37771; MUID:93252825; PMID:8486643
A;Accession: S37771
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1848 <BIR>
A;Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;48-80/Domain: ankyrin repeat homology <AN01>
F;81-113/Domain: ankyrin repeat homology <AN02>
F;114-146/Domain: ankyrin repeat homology <AN03>
F;147-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;275-307/Domain: ankyrin repeat homology <AN08>
F;308-340/Domain: ankyrin repeat homology <AN09>
F;341-373/Domain: ankyrin repeat homology <AN10>
F;374-406/Domain: ankyrin repeat homology <AN11>
F;407-439/Domain: ankyrin repeat homology <AN12>
```

F;440-472/Domain: ankyrin repeat homology <AN13>
F;473-505/Domain: ankyrin repeat homology <AN14>
F;506-538/Domain: ankyrin repeat homology <AN15>
F;539-571/Domain: ankyrin repeat homology <AN16>
F;572-604/Domain: ankyrin repeat homology <AN17>
F;605-637/Domain: ankyrin repeat homology <AN18>
F;638-670/Domain: ankyrin repeat homology <AN19>
F;671-703/Domain: ankyrin repeat homology <AN20>
F;704-736/Domain: ankyrin repeat homology <AN21>
F;737-769/Domain: ankyrin repeat homology <AN22>
F;770-802/Domain: ankyrin repeat homology <AN23>

Query Match 14.6%; Score 839.5; DB 2; Length 1848;
Best Local Similarity 29.5%; Pred. No. 4.9e-44;
Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |      :|||
Db         93 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 152

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                | :|  :|   |: :| :  :          || | |||          |||
Db        153 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 203

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
              ::|  :::|          |||::  |  :   :                  ||||:|:
Db        204 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 253

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
              | :|:|||  || :  : | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        254 GNVIMVRLLLDRGAQIETRTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 313

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:           ||
Db        314 HMAAQGDHLDCVRLLLQYNAE----------ID-------------------------- 336

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                | | || ||   :    :: ::|| |||  | :
Db        337 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 376

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : ||  |  ||:
Db        377 TPLHIACKKNHIRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 436

Qy        396 ISLQGFTALQM----GNENVQQLLQEGISLGNSEAD-----------------RQLLE- 432
              :::  | | |    |:  | : | :  :  |::|                  : |||
Db        437 SNVKVETPLHMAARAGHTEVAKYLLQNKAKANAKADDQTPLHCAARIGHTGMVKLLLEN 496

Qy        433 -----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474
                              ||: | |:|   |  : | |   :|   |||| || | :|
Db        497 GASPNLATTAGHTPLHTAAREGHVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVR 554

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533
              : | ||:| |  :|  | || ||| |  : : :: :||:  |   : :  |  :|||| |
Db        555 LAELLLEHDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 613

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592
              | : : |: : |||:|      ::  | ||| |  ::| |::  ||
Db        614 AKQNQIEVARSLLQYGGSANAESVQGVTPLHLAAQEGHTEMVALL--------------- 658

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652
                ||   | |  :  |   ||||| :   :: ||: |::||  |:|  :  |  |||
Db        659 -----LSKQANGNLGNKSG--LTPLHLVSQEGHVPVADVLIKHGVTVDATTRMGYTPLHV 711

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712
              |: ||:: :   |::: | |||  |  ::|||:|||:| | :  ||| :|| |   :  |
Db        712 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNG 771

Qy        713 QTPLDL------VSADDVSALLT 729
              |||  :      :|  ||  ::|
Db        772 TTPLAIAKRLGYISVTDVLKVVT 794
```

RESULT 11
I49502
ankyrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 13-Aug-1999
C;Accession: I49502
R;White, R.A.; Birkenmeier, C.S.; Peters, L.L.; Barker, J.E.; Lux, S.E.
Mamm. Genome 3, 281-285, 1992
A;Title: Murine erythrocyte ankyrin cDNA: Highly conserved regions of the regulatory doma
A;Reference number: I49502; MUID:92345717; PMID:1386265
A;Accession: I49502
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1862 <RES>
A;Cross-references: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C;Genetics:
A;Gene: Ank-1
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;40-72/Domain: ankyrin repeat homology <AN01>
F;73-105/Domain: ankyrin repeat homology <AN02>
F;106-138/Domain: ankyrin repeat homology <AN03>
F;139-167/Domain: ankyrin repeat homology <AN04>
F;168-200/Domain: ankyrin repeat homology <AN05>
F;201-233/Domain: ankyrin repeat homology <AN06>
F;234-266/Domain: ankyrin repeat homology <AN07>
F;267-299/Domain: ankyrin repeat homology <AN08>
F;300-332/Domain: ankyrin repeat homology <AN09>
F;333-365/Domain: ankyrin repeat homology <AN10>
F;366-398/Domain: ankyrin repeat homology <AN11>
F;399-431/Domain: ankyrin repeat homology <AN12>
F;432-464/Domain: ankyrin repeat homology <AN13>
F;465-497/Domain: ankyrin repeat homology <AN14>
F;498-530/Domain: ankyrin repeat homology <AN15>
F;531-563/Domain: ankyrin repeat homology <AN16>
F;564-596/Domain: ankyrin repeat homology <AN17>
F;597-629/Domain: ankyrin repeat homology <AN18>
F;630-662/Domain: ankyrin repeat homology <AN19>
F;663-695/Domain: ankyrin repeat homology <AN20>
F;696-728/Domain: ankyrin repeat homology <AN21>
F;729-761/Domain: ankyrin repeat homology <AN22>
F;762-794/Domain: ankyrin repeat homology <AN23>

Query Match 14.5%; Score 837.5; DB 2; Length 1862;
Best Local Similarity 29.5%; Pred. No. 6.7e-44;
Matches 237; Conservative 127; Mismatches 262; Indels 177; Gaps 20;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||  |:  ||:| |:   |  ||: |    | |||  || :||: |      :|||
Db         85 GQDEVVRELVNYGANVNAQSQKGFTPLYMAAQENHLEVVKFLLENGANQNVATEDGFTPL 144

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTI---------RNTDGRTALDLADPSAKAVLTGEYKKD 113
                | :|  :|   |: :| :  :          || | |||          |||
Db        145 AVALQQGHENVVAHLINYGTKGKVRLPALHIAARNDDTRTA---------AVLLQNDPNP 195

Qy        114 ELLESARSGNEEKMMALLTPLNVNCHASDGRKS-----------------TPLHLAAGY 155
              ::|  :::|          |||::  |  :   :                  ||||:|:
Db        196 DVL--SKTG--------FTPLHIAAHYENLNVAQLLLNRGASVNFTPQNGITPLHIASRR 245

Qy        156 NRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPL 215
              | :|:|||  || :  : | :| ||| |   ||  ::|:|: ||| : |      :|:
Db        246 GNVIMVRLLLDRGAQIETRTKDELTPLHCAARNGHVRISEILLDHGAPIQAKTKNGLSPI 305

Qy        216 HEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAARE 275
              | ||  : ::   ||| | |:           ||
Db        306 HMAAQGDHLDCVRLLLQYNAE----------ID-------------------------- 328

Qy        276 ADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335
              |:|                | | || ||   :    :: ::|| |||  | :
Db        329 -DIT---------------LDHLTPLHVAA---HCGHHRVAKVLLDKGAKPNSRALNGF 368

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395
              ||||:| :| |  |:|:::|  | ::|:   | | || |:: |||   : ||  |  ||:
```

```
Db        369 TPLHIACKKNHIRVMELLLKTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNV 428

Qy        396 ISLQGFTALQM----GNENVQQLLQEGISLGNSEAD-----------------RQLLE- 432

Db        429 SNVKVETPLHMAARAGHTEVAKYLLQNKAKANAKAKDDQTPLHCAARIGHTGMVKLLLEN 488

Qy        433 -----------------AAKAGDVETVKKLCTVQ-SVNCRDIEGRQSTPLHFAAGYNRVS 474

Db        489 GASPNLATTAGHTPLHTAAREGHVDTALALLEKEASQACMTKKG--FTPLHVAAKYGKVR 546

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWK-FTPLHEA 533

Db        547 LAELLLEHDAHPNAAGKNGLTPLHVAVHHNNLDIVKLLLPRGGSPH-SPAWNGYTPLHIA 605

Qy        534 AAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCL 592

Db        606 AKQNQIEVARSLLQYGGSANAESVQGVTPLHLAAQEGHTEMVALL---------------- 650

Qy        593 ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHN 652

Db        651 -----LSKQANGNLGNKSG--LTPLHLVSQEGHVLVADVLIKHGVTVDATTRMGYTPLHV 703

Qy        653 AASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712

Db        704 ASHYGNIKLVKFLLQHQADVNAKTKLGYSPLHQAAQQGHTDIVTLLLKNGASPNEVSSNG 763

Qy        713 QTPLDL------VSADDVSALLT 729

Db        764 TTPLAIAKRLGYISVTDVLKVVT 786
```

RESULT 12
T13940
ankyrin - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C;Accession: T13940
R;Dubreuil, R.R.; Yu, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994
A;Title: Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Drosoph
A;Reference number: Z17820; MUID:95024098; PMID:7937942
A;Accession: T13940
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1549 <DUB>
A;Cross-references: EMBL:L35601; NID:g557083; PID:g557084; PIDN:AAC37208.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0011747

Query Match 14.5%; Score 834; DB 2; Length 1549;
Best Local Similarity 27.4%; Pred. No. 8.3e-44;
Matches 298; Conservative 165; Mismatches 470; Indels 154; Gaps 32;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         84 GYVDICCELLRRGIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPL 143

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-----DPSAKAVLTGEYK---KDE 114

Db        144 YMAAQENHDNCCRTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLENDVRGKVRLP 203

Qy        115 LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174

Db        204 ALHIAAKKNDVNAAKLLLQHDPNADIVSKSGFTPLHIAAHYGNVDIATLLLNNKADVNYV 263

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234

Db        264 AKHNITPLHVACKWGKLSLCTLLLCRGAKIDAATRDGLTPLHCASRSGHVEVIKHLLQQN 323

Qy        235 ADPTLLNCHN-KSAIDLAPTPQLKERLAYEPKGHSLLQAAREADVTRIKKHLSLE----- 288

Db        324 A-PILTKTKNGLSALHMAAQGE------HDEAAHLLLDNKAPVDEVTVDYLTALHVAAHC 376
```

```
Qy        289 --------MVNFK-HPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFL 335

Db        377 GHVKVAKLLLDYKANPNARALNGFTPLHIACKK---NRIKMVELLIKHGANIGATTESGL 433

Qy        336 TPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNI 395

Db        434 TPLHVASFMGCINIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLR-SAKVDA 492

Qy        396 ISLQGFTAL----QMGNEN-VQQLLQEGISLGNSEADR--QLLEAAKAGDVETVKKLCTV 448

Db        493 IVREGQTPLHVASRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENIVQVLLEN 552

Qy        449 QSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEV 508

Db        553 GAEN-NAVTKKGFTPLHLACKYGKQNVVQILLQNGASIDFQGKNDVTPLHVATHYNNPSI 611

Qy        509 AELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VK 567

Db        612 VELLLKNGSSPNLCARNGQCAIHIACKKNYLEIAMQLLQHGADVNIISKSGFSPLHLAAQ 671

Qy        568 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLE 627

Db        672 GGNVDMVQLLLEYGVISAAAKNGL----------------------TPLHVAAQEGHVL 708

Qy        628 VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA 687

Db        709 VSQILLEHGANISERTRNGYTPLHMAAHYGHLDLVKFFIENDADIEMSSNIGYTPLHQAA 768

Qy        688 QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLN 747

Db        769 QQGHIMIINLLLRHKANPNALTKDGNTALHIASN------LGYVTVMESLKIVTSTSVIN 822

Qy        748 G----------VRSPGATADALSSGPSSPSSLSAAS-------SLDNLSGSFSELSSVVS 790

Db        823 SNIGAIEEKLKVMTPELMQETLLSDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFD 882

Qy        791 SSGTE----GASSLEKKEVPGVDFSITQFVRNLGLEHLMD--IFEREQITLDVLVEMGHK 844

Db        883 TTNTDHDLTDVSVLNKKEILPNEMSCIELTE---IGHKPDNVVIARSQVHLGFLVSF--- 936

Qy        845 ELKEIGINAYGHRHKLIKGVE------------RLISGQQGLNPYLTLNTSGSGTILID 891

Db        937 LVDARGGSMRGYRHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSRILE 996

Qy        892 LSPDDKEFQS---VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRK 948

Db        997 MSPVDGMFLSPITLEVPHYGTLRK----------NEREII-ILRSDNGESWREH-NLYK 1043

Qy        949 EVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYG 1008

Db       1044 DIIGEDINQTEE---FHSDRIVR-IVTQNVPHFFAVV-------------SRVRQEVHV 1085

Qy       1009 IG--GGT 1013

Db       1086 IGPDGGT 1092
```

RESULT 13
T15347
ankyrin-related unc-44 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Aug-2002
C;Accession: T15347; T15346; T15344; T15345; A57282; B57282; C57282
R;Gattung, S.
submitted to the EMBL Data Library, February 1996
A;Description: The sequence of C. elegans cosmid B0350.
A;Reference number: Z18332
A;Accession: T15347
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2039 <GAT>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208873; PIDN:AAA93443.1
A;Accession: T15346

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1000,'SKLQHRT',1002-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVA
32,'S',2034-2035,'GSPTRRSVEPEEHRHSQHEDHEGST' <GA2>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208875; PIDN:AAA93445.1
A;Accession: T15344
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1718,'KW',1903-1905,'NRLADESSPS',1916-1917,'QRSTIVAESTSEQVPE',1934-1935,'E
<GA3>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208874; PIDN:AAA93444.1
A;Accession: T15345
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1718,'KWEELNRL',1727,'D','ES',1942,'PSPAQRS',1950,'IVAES',1956-1957,'EQVPE
PTRRSVEPEEHRHSQHEDHEGST' <GA4>
A;Cross-references: EMBL:U50071; NID:g1208871; PID:g1208876; PIDN:AAA93446.1
R;Otsuka, A.J.; Franco, R.; Yang, B.; Shim, K.H.; Tang, L.Z.; Zhang, Y.Y.; Boontrakulpoo
J. Cell Biol. 129, 1081-1092, 1995
A;Title: An ankyrin-related gene (unc-44) is necessary for proper axonal guidance in Cae
A;Reference number: A57282; MUID:95263663; PMID:7744957
A;Accession: A57282
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-852,'GGG',856-1000,'SKLQHRT',1002-1319,'IG',1322-1595,'DA',1598-1718,'KWEE
,'SHRED',2007-2008,'TI',2011,2017,'TT',2020-2022,'SHIS' <OTS>
A;Cross-references: GB:U21734; NID:g790607; PIDN:AAA85854.1; PID:g790608
A;Accession: B57282
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 831-852,'GGG',856-1319,'IG',1322-1595,'DA',1598-1718,'KWEELNRL',1727,1799,'E
V',1945-1947,'VT',1950,'SH',1975,'SESP',1980-1981,'SPTRRSVEPEEHRHS',1984-1985,'EDHEGS',1
A;Cross-references: GB:U21731
A;Accession: C57282
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 194,'F',196,'I',198,'DC',201,'G',409,'AV',412,'Q',414,824,'S',826,'ER',829,'
4,'TIV',1828,'ESTS',1833,'QV',1836,'E',1934-1935,'EQS',1939,'ESES',1944,'REDDGTIVTT',194
A;Cross-references: GB:U21732; NID:g790603; PIDN:AAA85853.1; PID:g790604
C;Genetics:
A;Gene: CESP:unc-44
A;Introns: 97/3; 262/2; 535/1; 749/1; 786/3; 883/3; 1001/2; 1045/3; 1660/1; 1719/2; 1979
C;Superfamily: ankyrin; ankyrin repeat homology
F;164-192/Domain: ankyrin repeat homology <AN04>
F;358-390/Domain: ankyrin repeat homology <AN1>
F;391-423/Domain: ankyrin repeat homology <AN11>

  Query Match           14.1%;  Score 814.5;  DB 2;  Length 2039;
  Best Local Similarity  26.1%;  Pred. No. 2.1e-42;
  Matches  301;  Conservative  156;  Mismatches  418;  Indels  279;  Gaps   36;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :||  |::  | | |     |   || |   | : :| :|: :||: | :    :|||
Db         77 GHSEVVRELIKRQAQVDAATRKGNTALHIASLAGQSLIVTILVENGANVNVQSVNGFTPL 136

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
              : || :   :|   ||:|||   :   || | | :|                    : |
Db        137 YMAAQENHEEVVKYLLKHGANQALSTEDGFTPLAVA--------------------LQQG 176

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181
              ::  :  ||       : | |:   | ||:|  :      ||||:  :      |    |
Db        177 HDRVVAVLLE------NDSKGKVRLPALHIAAKKDDTTAATLLLQNEHNPDVTSKSGFTP 230

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
              || |  |||  | :||:: || ||      :||| |     |  : :|||| ||   :::
Db        231 LHIAAHYGHENVGQLLLEKGANVNYQARHNISPLHVATKWGRTNMANLLLSRGA---IID 287

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301
              |  :                                                      |
Db        288 SRTKDLL-------------------------------------------------TP 296

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
              ||||| | :    |: :||: :|| |: |||  | |||:|::  | |    :: | | |:

Db        297 LHCAARSGH---DQVVDLLVVQGAPISAKTKNGLAPLHMAAQGDHVDAARTLLYHRAPVD 353

Qy        362 --ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN--------- 410
                :| |  | || ||:|||::  :|||    |||  :| ||| | :  :
Db        354 DVTVDYL--TPLHVAAHCGHVRVAKLLLDRSADPNSRALNGFTPLHIACKKNRIKVVELL 411

Qy        411 ------VQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIE 457
                   ::   : |::       :|       ||:     |||||
Db        412 LKYRAAIEATTESGLTPLHVAAFMGAINIVIYLLQQGANPDVETV--------------- 456

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517
              |  |||| ||  |:  ||  |:::|| | |: :    ||| |   |: ::  ||:: ||
Db        457 -RGETPLHLAARANQTDVVRVLIRNGAKVDAQARELQTPLHIASRLGNTDIVILLLQAGA 515

Qy        518 VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDL 576
              |       ::||| || :|: |:  :|| | || |    : | ||| |   | |: ::  |
Db        516 NSNATTRDNYSPLHIAAKEGQEEVAGILLDHNADKTLLTKKGFTPLHLASKYGNLEVVRL 575

Qy        577 L--RGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLL 633
              |  ||                    | :|::  ||||:|| ||| :||  ||
Db        576 LLERGTPV------------------------DIEGKNQVTPLHVAAHYNNDKVAMLLL 610

Qy        634 QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQ 693
              ::||    |  | |  ||| ||    :::|: |::: |  ||  :  |||||  :||:|  :
Db        611 ENGASAKAAAKNGYTPLHIAAKKNQMEIASTLLQFKADPNAKSRAGFTPLHLSAQEGHKE 670

Qy        694 LCALLLAHGADPTLKNQEGQTPLDLVSADD-------------VSALLTAAMPPSALPS 739
              :  ||: :|:|    |   | | : :|                :::   |   | : :
Db        671 ISGLLIENGSDVGAKANNGLTAMHLCAQEDHVPVAQILYNNGAEINSKTNAGYTPLHV-A 729

Qy        740 CYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
              |:  | || |:      ||      :    :|   :|    :|       |          |
Db        730 CHFGQ-LNMVKFLVENGADVGEKTRASYTPLHQAAQQGHNNCVRYLLE----------NG 778

Qy        797 ASSLEKKEVPGVDFSITQFVRNLG----LEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852
              ||  |:              ||  |    :| |  : |   ||    |:  :|      :|
Db        779 ASPNEQTATGQTPLSIAQ---RLGYVSVVETLRTVTETTVITETTTVDERYKPQNPEAMN 835

Qy        853 ----------AYGHRHKLI-----------------KGVERLISGQQGLNPYLTLNTSG 884
                         |: :                   ||  : :|:| |      |   :|
Db        836 ETMFSESEDEGQAAEHEAVAAHEKDFSDNLTQGLQDSTGVHMIHTGEQLLQRSQELENGG 895

Qy        885 ------SGTILIDLSPDDKEFQSVEEEMQS----TVREHRDG--GHAGGIFNRYNILKIQ 932
                   ||     :|| :|||  :     |     |       |    || |  ::    :
Db        896 AIPKINSG----GMSP-EKEFAKIAPVATSSPIATSNSQSFGIAPRAGSISGQFQQQPLH 950

Qy        933 KVCNKKLWERYTHRRKE--VSEENHNHANERMLFHGS----PFVNAIIHKGF------DE 980
                :  |    | :   ::  |:::    || :|      |  : :    |       |
Db        951 GAGPEDNLEELVRRAQNHPINAGNYDNGGVAMLENGHADNVPIGHHVTQPSFLISFLVDA 1010

Qy        981 RHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKS 1040
              |   : |   :|:        |::|          |:                : ||  | |
Db       1011 RGGAMRGCRHSGVRIIVPPRKASQ----------PIR----------VTCRY-LRKD 1046

Qy       1041 FLQFSAMKMAHSPP 1054
                 |:|| ||
Db       1047 -------KLAHPPP 1053


RESULT 14
T42691
hypothetical protein DKFZp434D2328.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T42691
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, November 1999
A;Reference number: Z22230
A;Accession: T42691
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-791 <AAA>
A;Cross-references: EMBL:AL133087
```

A;Experimental source: adult testis; clone DKFZp434D2328
C;Genetics:
A;Note: DKFZp434D2328.1

Query Match 10.9%; Score 628; DB 2; Length 791;
Best Local Similarity 27.7%; Pred. No. 2.5e-31;
Matches 232; Conservative 119; Mismatches 313; Indels 174; Gaps 28;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |: :||::||  |  :   :  |   || ||  |   ||| |: :||: |  :|  :|||
Db         14 GQINVVKHLLNLGVEIDEINVYGNTALHIACYNGQDAVVNELIDYGANVNQPNNNGFTPL 73

Qy         63 HEAAIKGKIDVCI-VLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELL----- 116
              | ||      :|: :|: :||:  |:: ||:: | :       : | : : : |
Db         74 HFAAASTHGALCLELLVNNGADVNIQSKDGKSPLHM------TAVHGRFTRSQTLIQNGG 127

Qy        117 -------------ESARSGNEEKMMALLT-------------PLN---VNCHA----- 140
                          :|| |:|  :  :|                ||:  :| |:
Db        128 EIDCVDKDGNTPLHVAARYGHELLINTLITSGADTAKCGIHSMFPLHLAALNAHSDCCRK 187

Qy        141 --SDGRK-----------------STP-------LHLAAGYNRVKIVQLLLQHGADVH 172
                | |:|                 ||       || ||    |: ::||   ||| |
Db        188 LLSSGQKYSIVSLFSNEHVLSAGFEIDTPDKFGRTCLHAAAAGGNVECIKLLQSSGADFH 247

Qy        173 AKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLS 232
               |||    ||| | :  |:   | ||  || ||  | | |  | || ||: :
Db        248 KKDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDDWGRTALHYAAASD---------- 297

Qy        233 YGADPTLL-NCHNKSAIDLAPTPQLKER---LAYEFKGHSLLQAARE-ADVTRIKKHLSLE 288
              : |:| | |: |  :|    :|||:   |  ||    ||
Db        298 MDRNKTILGNAHDNSE-ELERARELKEKEATLCLEF----LLQ----------------- 335

Qy        289 MVNFKHPQTHE----TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT--PLHVAS 342
                |  :|   :     ::| |||  :   :|  |||| :  :  |::    |  |||:|:
Db        336 --NDANPSIRDKEGYNSIHYAAAYGH---RQCLELLLERTNSGFEESDSGATKSPLHLAA 390

Qy        343 EKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFT 402
               |:  :||:::    ::  |  |:|:|  ||: || :    |::           || :
Db        391 YNGHHQALEVLLQSLVDLDIRDEKGRTALDLAAFKGHTECVEALIN----------QGAS 440

Qy        403 ALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQST 462
                  :  |   :  |::  | ||| |                ::|: :| :|:  |
Db        441 IFVKDNVTKRTPLHASVINGHTLCLRLLLEIAD----------NPEAVDVKDAKGQ--T 487

Qy        463 PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 522
              ||  |  |  :  |  ||:  |:|   |  |   ||   || |  ::|::   :
Db        488 PLMLAVAYGHIDAVSLLLEKEANVDTVDILGCTALHRGIMTGHEECVQMLLEQEVSILCK 547

Qy        523 DLWKFTPLHEAAAKGKYEICKLLLQHG---ADPTKKNRDGNTPLD--LVKDGDTDIQDLL 577
              |   |||| |||:|      |||       |   |:  | |||        :  |: ||
Db        548 DSRGRTPLHYAAARGHATWLSELLQMALSEEDCCFKDNQGYTPLHWACYNGNENCIEVLL 607

Qy        578 R--------GD-------AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAG 622
                      |:       | : |      | : :          |:||| :||  |||| ||
Db        608 EQKCFRKFIGNPFTPLHCAIINDHG--NCASLLLGAIDSSIVSCRDDKGR--TPLHAAAF 663

Qy        623 YNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLI-KYNACVNATDKWAFT 681
               :::|  : ||:| | |||  |   |  || |       :|:    |  | |  ||  |
Db        664 ADHVECLQLLLRHSAPVNAVDNSGKTALMMAAENGQAGAVDILVNSAQADLTVKDKDLNT 723

Qy        682 PLHEAAQKGRTQLCALLLAHGADPTL---KNQEGQTPLDLVSADDVSALLTAAMPPSA 736
              ||| |  ||  :   |:|    | :|   ||   |||| : : : :  ::   :  |
Db        724 PLHLACSKGHEKCALLILDKIQDESLINEKNNALQTPLHVAARNGLKVVVEELLAKGA 781
```

RESULT 15
S30355
alpha-latroinsectotoxin precursor - black widow spider (fragment)
C;Species: Latrodectus mactans tredecimguttatus (black widow spider)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Sep-1999
C;Accession: S30355; S29864
R;Kiyatkin, N.; Dulubova, I.; Grishin, E.
Eur. J. Biochem. 213, 121-127, 1993

A;Title: Cloning and structural analysis of alpha-latroinsectotoxin cDNA. Abundance of a
A;Reference number: S30355; MUID:93238678; PMID:8477689
A;Accession: S30355
A;Molecule type: mRNA
A;Residues: 1-1411 <KIY>
A;Cross-references: EMBL:Z14086; NID:g9536; PIDN:CAA78464.1; PID:g9537
A;Note: it is uncertain whether Met-26 is the initiator or whether translation is initia
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: presynaptic neurotoxin
F;533-565/Domain: ankyrin repeat homology <AN1>
F;1001-1032/Domain: ankyrin repeat homology <AN2>
F;1146-1178/Domain: ankyrin repeat homology <AN3>

Query Match 9.5%; Score 545.5; DB 2; Length 1411;
Best Local Similarity 23.7%; Pred. No. 9.4e-26;
Matches 239; Conservative 136; Mismatches 378; Indels 257; Gaps 33;

```
Qy         11 LLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLR------HGADPNARDNWNYTPLHE 64
              |:||||:|    :  |  :| | | |: :|  |||          || |      |||||
Db        486 LIQNGANVSETFELGRGAIHAAASAGNYDVGELLLNKDINLLEKADKN-----GYTPLHI 540

Qy         65 AAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNE 124
              ||   | |  : |: : |:  :     ||  ||  |                | :||
Db        541 AADSNKNDFVMFLIGNNADVNV-----RTKSDLFTP--------------LHLAARRDLT 581

Qy        125 EKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 184
              :    |:    :: :|  |     |||||:           |:     | :: | |  | |||
Db        582 DVTQTLIDITEIDLNAQDKSGFTPLHLSISSTSETAAILIRNTNAVINIKSKVGLTPLHL 641

Qy        185 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 244
              |    :  |::||   || :|  |    |||| ||    :|:   ||            |
Db        642 ATLQNNLSVSKLLAGKGAYLNDGDANGMTPLHYAAMTGNLEMVDFLL-----------N 689

Qy        245 KSAIDLAPTPQLKE----RLAYEFKGHSLLQ---------------------------- 271
              :  |::    : |:      ||  || : : :
Db        690 QQYININAATKEKKWTPLHLAILFKKNDVAERLLSDENLNIRLETNGGINPLHLASATGN 749

Qy        272 -------AAREADVTRI-KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK 323
                    |: |||||:  |  |                ||| |  : ::|   |: |
Db        750 KQLVIELLAKNADVTRLTSKGFS-------------ALHLGIIG---KNEEIPFFLVEK 792

Qy        324 GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTC 383
              |||:|:||   :|||| |:     ::  :::   | : | |   |  :| | |  |||:
Db        793 GANVNDKTNSGVTPLHFAAGLGKANIFRLLLSRGADIKAEDINSQMPIHEAVSNGHLEIV 852

Qy        384 RLLLSYGCDPNIISLQGFTALQMGNENVQQLLQE-GISLGNSEADRQLLEAAKAGDVETV 442
              |:|:   ||:::              ||: :  |    |   :  : : :  : |
Db        853 RILIEK--DPSLM-------------NVKNIRNEYPFYLAVEKRYKDIFDYFVSKD---- 893

Qy        443 KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKD------------ 490
                  |  :::   :| ||  :    : ||::|:|:||:   |:
Db        894 --------ANVNEVDHNGNTLLHLFSSTGELEVVQFLMQNGANFRLKNNERKTFFDLAIE 945

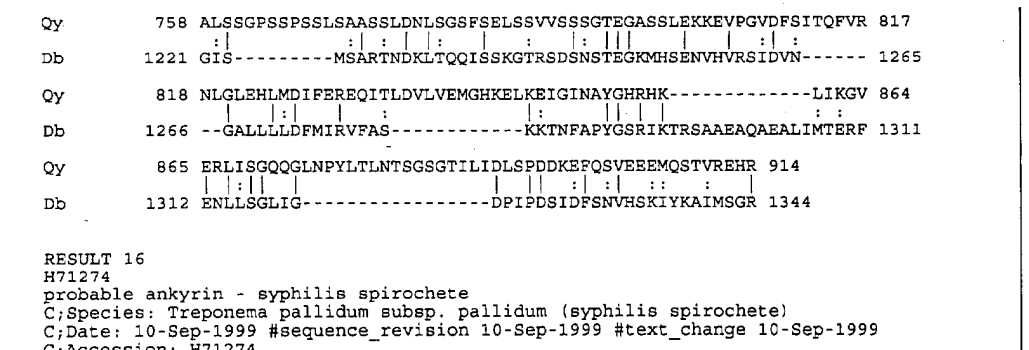
Qy        491 ----------------------KGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFT 528
                                   :|  :  |  |    |:  |::       :| ::
Db        946 NGRLNIVAFAVEKNKVNLQAAHRGKTILYHAICDSAKYDKIEIVKYFIEKLNESEC---N 1002

Qy        529 PLHEAAAKGKYEICKLLLQ-HGADPTKKNRDGN-TPLDLVKDG-------DTDIQDLLRG 579
              |||||||    :: |  :|  | :| : | :   :|  :  |       |  | | |
Db       1003 PLHEAAAYAHLDLVKYFVQERGINPAEFNEENQASPFCITIHGAPCGYSLDCDTPDRLEV 1062

Qy        580 DAALLDAAKKGCLARVKKLSSPD-NVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGAD 638
              | |                || |  | | |   :||: :|   | : :  ||:  |||
Db       1063 VEYLSD-------------KIPDINGKC-DVQ--ENTPITVAIFANKVSILNYLVGIGAD 1106

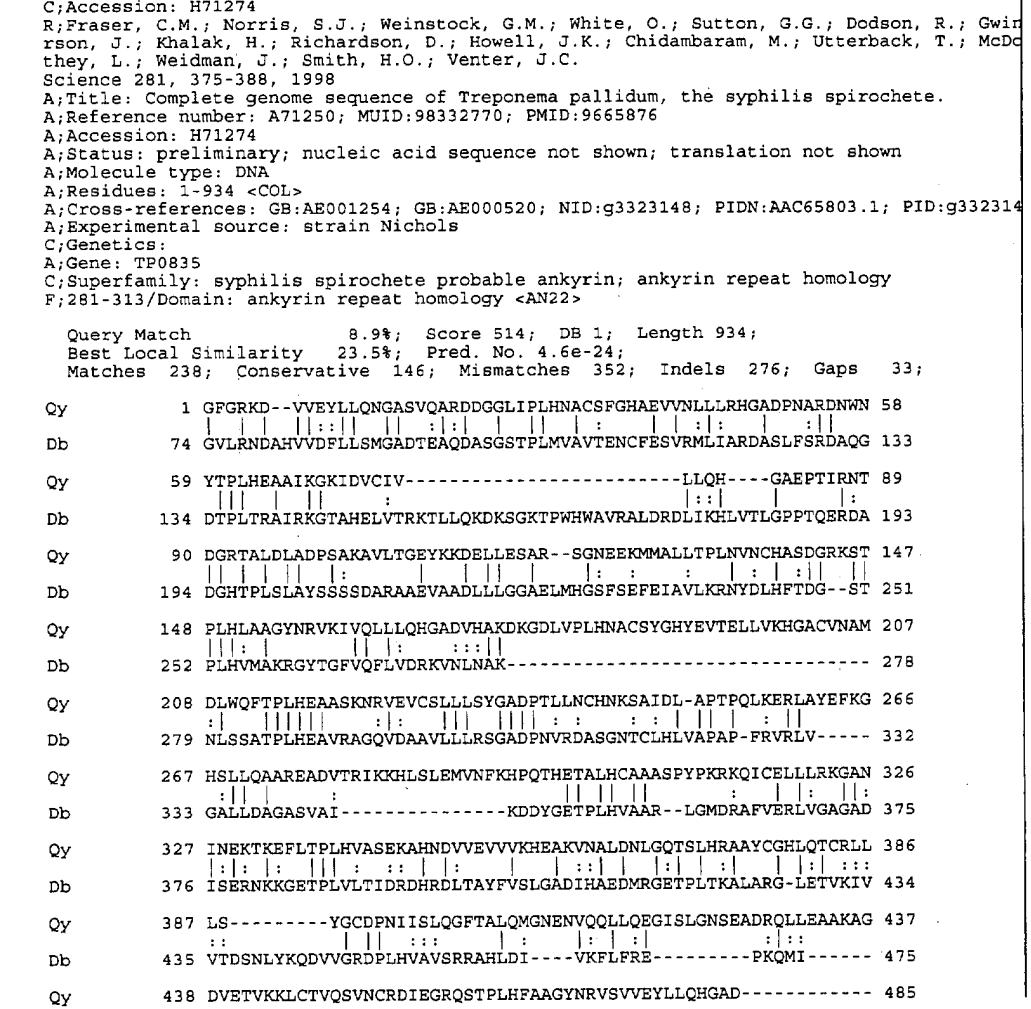
Qy        639 VNAQDKGGLIPLHNAASYGHVDVAALLIK-YNACVNATDKWAFTPLHEAAQKGRTQLCAL 697
              | |  |   ||: || |  ::   ||: :  :|  :|  || || ||:     :
Db       1107 PNQQVDGD-PPLYIAARQGRFEIVRCLIEVHKVDINTRNKERFTALHAAARNDFMDVVKY 1165

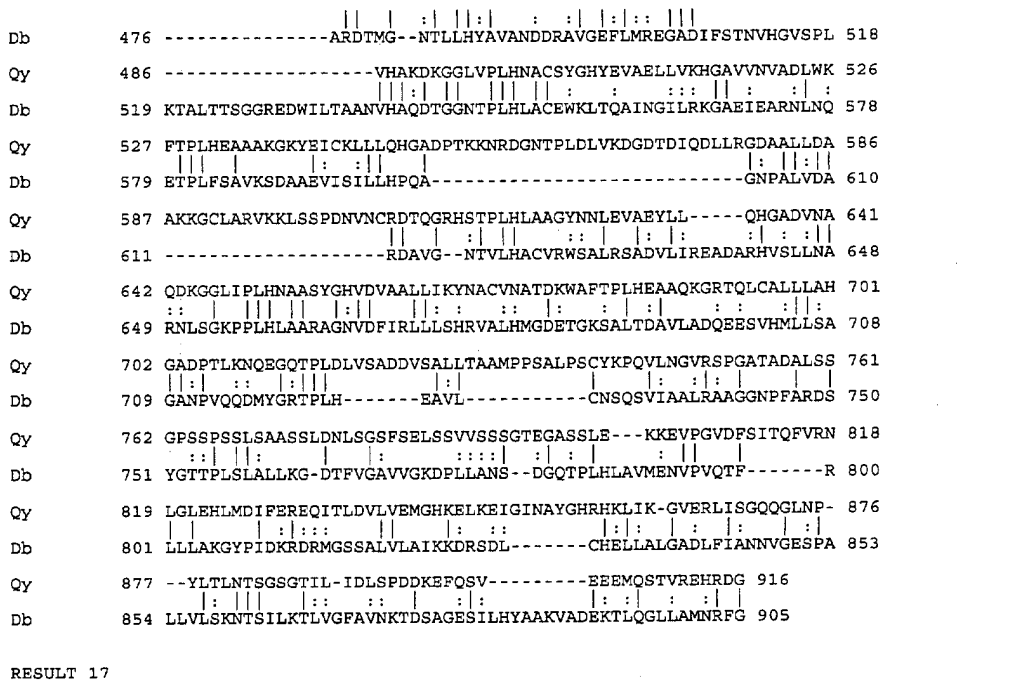
Qy        698 LLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATAD 757
              |:  |||   |  :   |:| :: :   | | ::      | | :  |  | : | | |
Db       1166 LVRQGADVNAKGIDDLRPID-IAGEKAKAYLQSS---RFLRSGHSFQ-SNEIDSFGNTIH 1220
```



RESULT 16
H71274
probable ankyrin - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: H71274
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: H71274
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-934 <COL>
A;Cross-references: GB:AE001254; GB:AE000520; NID:g3323148; PIDN:AAC65803.1; PID:g332314
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0835
C;Superfamily: syphilis spirochete probable ankyrin; ankyrin repeat homology
F;281-313/Domain: ankyrin repeat homology <AN22>

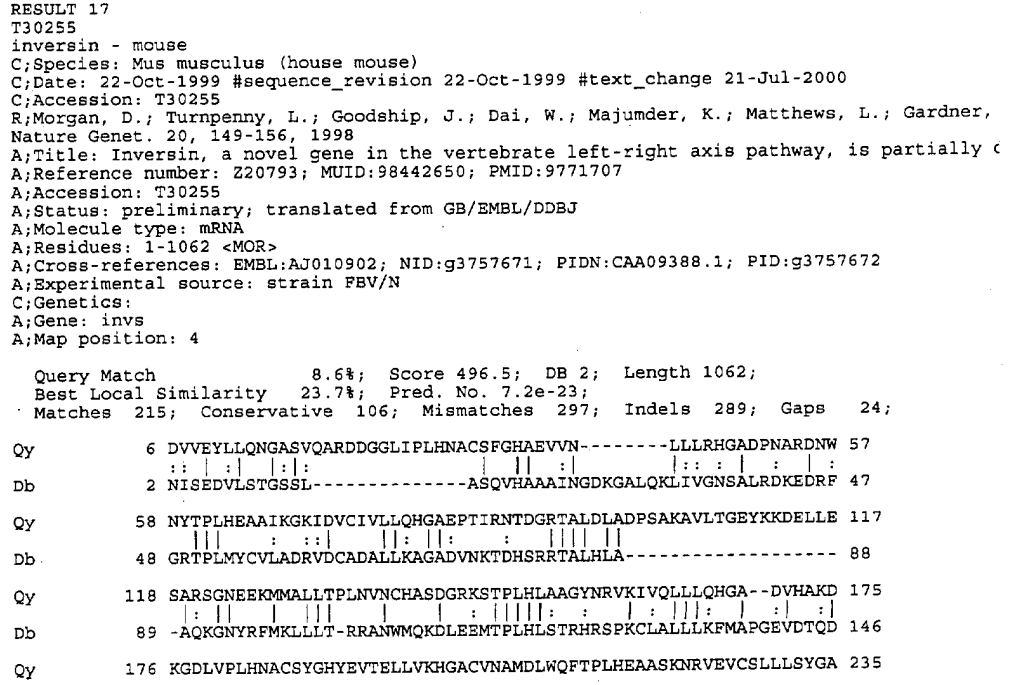
Query Match 8.9%; Score 514; DB 1; Length 934;
Best Local Similarity 23.5%; Pred. No. 4.6e-24;
Matches 238; Conservative 146; Mismatches 352; Indels 276; Gaps 33;

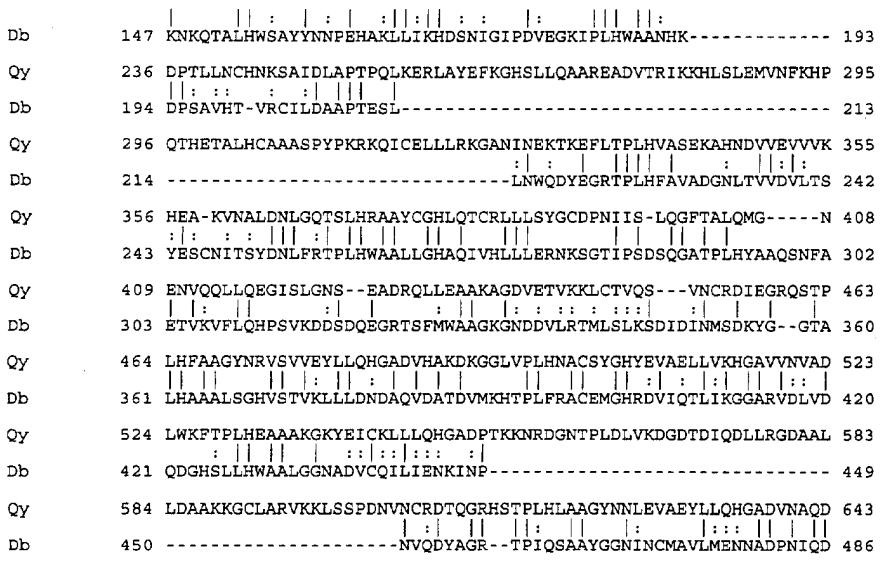




RESULT 17
T30255
inversin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C;Accession: T30255
R;Morgan, D.; Turnpenny, L.; Goodship, J.; Dai, W.; Majumder, K.; Matthews, L.; Gardner,
Nature Genet. 20, 149-156, 1998
A;Title: Inversin, a novel gene in the vertebrate left-right axis pathway, is partially c
A;Reference number: Z20793; MUID:98442650; PMID:9771707
A;Accession: T30255
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1062 <MOR>
A;Cross-references: EMBL:AJ010902; NID:g3757671; PIDN:CAA09388.1; PID:g3757672
A;Experimental source: strain FBV/N
C;Genetics:
A;Gene: invs
A;Map position: 4

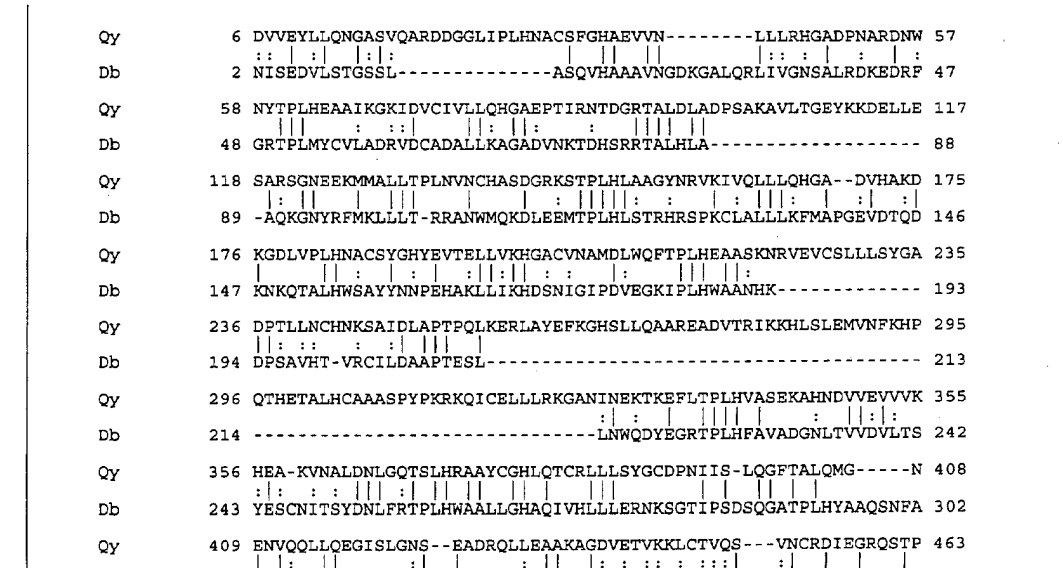
Query Match 8.6%; Score 496.5; DB 2; Length 1062;
Best Local Similarity 23.7%; Pred. No. 7.2e-23;
Matches 215; Conservative 106; Mismatches 297; Indels 289; Gaps 24;





RESULT 18
T14151
Inv protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14151
R;Yokoyama, T.; Mochizuki, T.
submitted to the EMBL Data Library, November 1997
A;Reference number: Z17888
A;Accession: T14151
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1062 <YOK>
A;Cross-references: EMBL:AF034860; NID:g3560522; PID:g3560523; PIDN:AAC34976.1
C;Genetics:
A;Gene: inv
A;Map position: 4

Query Match 8.5%; Score 491.5; DB 2; Length 1062;
Best Local Similarity 23.9%; Pred. No. 1.5e-22;
Matches 217; Conservative 103; Mismatches 298; Indels 289; Gaps 25;



RESULT 19
S11527
alpha-latrotoxin precursor - black widow spider
C;Species: Latrodectus mactans tredecimguttatus (black widow spider)
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Sep-1999
C;Accession: S11527; JN0313
R;Kiyatkin, N.I.; Dulubova, I.E.; Chekhovskaya, I.A.; Grishin, E.V.

```
FEBS Lett. 270, 127-131, 1990
A;Title: Cloning and structure of cDNA encoding alpha-latrotoxin from black widow spider
A;Reference number: S11527; MUID:91031994; PMID:1977615
A;Accession: S11527
A;Molecule type: mRNA
A;Residues: 1-1401 <KIY>
A;Cross-references: EMBL:X55009; NID:g9568; PIDN:CAA38753.1; PID:g9569
A;Note: it is uncertain whether Met-1, Met-8 or Met-18 is the initiator
R;Volkova, T.M.; Galkina, T.G.; Kudelin, A.B.; Grishin, E.V.
Bioorg. Khim. 17, 437-441, 1991
A;Title: Tryptic fragments structure of the black widow spider venom neurotoxin.
A;Reference number: JN0313; MUID:91362695; PMID:1888339
A;Note: English title and abstract; article in Russian
A;Accession: JN0313
A;Molecule type: protein
A;Residues: 21-33,'X',35-40,'X',43-50;82-90,'X',92-99;107-113;114-126;146-151;171-178;19
84-792;881-891;906-913;914,1015-1021;1062-1065,'X',1067-1075;1080-1084;1087-1096;1183-11
A;Note: the sequence of the amino terminal peptide has been corrected with tryptic pepti
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: neurotoxin; venom
F;1-20/Domain: propeptide #status predicted <PRO>
F;21-1401/Product: alpha-latrotoxin #status experimental <MAT>
F;525-557/Domain: ankyrin repeat homology <AN7>
F;862-894/Domain: ankyrin repeat homology <AN3>
F;1005-1036/Domain: ankyrin repeat homology <AN21>

  Query Match             8.1%;  Score 464.5;  DB 2;  Length 1401;
  Best Local Similarity   23.2%;  Pred. No. 1.1e-20;
  Matches  175;  Conservative   94;  Mismatches  274;  Indels  211;  Gaps   19;

Qy         12 LQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLHEAAIKGK 70
Db        513 LAKSTELNQPDKKGYTPIHVAADSGNAGIVNLLIQRGVSINSKTYHFLQTPLHLAAQRGF 572

Qy         71 IDVCIVLLQHGAEPTI----RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEK 126
Db        573 VTTFQRLME---SPEININERDKDGFTP--------------------LHYAIRGGERI 608

Qy        127 MMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAKDKGDLVPLHNA 185
Db        609 LEAFLNQISIDVNAKSNTGLTPFHLAIIKNDWPVASTLLGSKKVDINAVDENNITALHYA 668

Qy        186 CSYGHYEVTE------------------------LLVKH-----------GACVNAMDL 209
Db        669 AILGYLETTKQLINLKEINANVVSSPGLLSALHYAILYKHDDVASFLMRSSNVNVNLKAL 728

Qy        210 WQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSL 269
Db        729 GGITPLHLAVIQGRKQILSLMFDIG---------------------------------- 753

Qy        270 LQAAREADVTRIKKHLSLEMVNFKHPQTHE--TALHCAAASPYPKRKQICELLLRKGANI 327
Db        754 -------------------VNIEQ-KTDEKYTPLHLAAMSKYP---ELIQILLDQGSNF 789

Qy        328 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 387
Db        790 EAKTNSGATPLHLATFKGKSQAALILLNNEVNWRDTDENGQMPIHGAAMTGLLDVAQAII 849

Qy        388 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 447
Db        850 S--IDATVV---------------------------------------------------- 856

Qy        448 VQSVNCRDIEGRQS-TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 506
Db        857 -------DIEDKNSDTPLNLAAQNSHIDVIKYFIDQGADINTRNKKGLAPLLAFSKKGNL 909

Qy        507 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 566
Db        910 DMVKYLFDKNANVYIADNDGMNFFYYAVQNGHLNIVKYAMSEKDKFEWSNTDNN------ 963

Qy        567 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHST--PLHLAAGYN 624
Db        964 RRDECPNEECAISHFAVCDAVQFDRIEIVKYFVG--------TLGNFAICGPLHQAARYG 1015

Qy        625 NLEVAEYLLQH---GADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFT 681
Db       1016 HLDIVKYLVEEEFLSVDGSKTD----TPLCYASENGHFTVVQYLVSNGAKVNHDCGNGMT 1071

Qy        682 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 715
Db       1072 AIDKAITKNHLQVVQFLAANGVDFRRKNSRGTTP 1105


RESULT 20
T19552
hypothetical protein C29E6.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19552
R;Dobson, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z19141
A;Accession: T19552
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1188 <WIL>
A;Cross-references: EMBL:Z72504; PIDN:CAA96603.1; GSPDB:GN00022; CESP:C29E6.2
A;Experimental source: clone C29E6
C;Genetics:
A;Gene: CESP:C29E6.2
A;Map position: 4
A;Introns: 147/2; 473/3; 500/1; 529/1; 584/2; 688/2; 839/3; 975/3; 1132/3

  Query Match             7.9%;  Score 456.5;  DB 2;  Length 1188;
  Best Local Similarity   22.8%;  Pred. No. 2.8e-20;
  Matches  201;  Conservative  130;  Mismatches  364;  Indels  185;  Gaps   24;

Qy          9 EYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAA-- 66
Db         50 EALLKAPLAVNAQDGDFMTPLHYAARYGNYDAVKLLLSKNALPNTKNREGDTPLHIASKY 109

Qy         67 IKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGN--E 124
Db        110 IYGYSDICSII---------------------------------DEDQADSARKYNTAT 135

Qy        125 EKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHN 184
Db        136 KKIINALVSENAEIDPVNKYQLTPLHYAAMKSNFSALHALIKLKADVDAEDDNKMTPLLL 195

Qy        185 ACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHN 244
Db        196 ACVHGSQEIIQELIKANSNVTKRDQRLNTVFHIVALRGEPEYLEMMMD------------ 243

Qy        245 KSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEM-VNFKHPQTHETALH 303
Db        244 ----------------------HDPVEAIKALNLFNNEKKTPLRMAVEGNHPETLKKILQ 281

Qy        304 CAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNAL 363
Db        282 M-------EKKNSCKWM--------DREKEL---IHFAAEKGFLEVLKALVEAGGNKNEL 323

Qy        364 DNLGQTSLHRAAYCGHLQTCRLLLSYGCDP-NIISLQGFTALQMGNENVQQLLQEGISLG 422
Db        324 NEVKAVPLHVAAQMNQLEVVSYLIEEEKDNIDVVDEQGLTPLMMAVTH------------ 371

Qy        423 NSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 482
Db        372 DSKKCVEYLIAKKANLTITDKD---------------ERTPVFIGAKFNALSSVEYILDH 416

Qy        483 GADVHAKDKGGLV--PLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
Db        417 LRKKNKETERSALKSPTRNTLRIVSEDVRR------TMVNMVDRDQNTPMHIVASNGYLE 470

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
Db        471 MMQLLQKHGASITQVNEDEETALHRAAIG---VRQLLEWDIRLL--------------- 511
```

```
Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
Db        512 ----LMKDEMG--NSALHLAARSGHDATTKVLLDNGADKEAKNSYQKTPLQVAVDSGKLE 565

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
Db        566 TCQRLVAKGAQIESSSD-TKTVLHTAAFYGNESIVRYFIAEGVTIDRRDEEGKTAFDIAC 624

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALS-SGPSSPSSLSAASSLDNLS 779
Db        625 ENDHKDVARAFLETDQWKNLMIP------------CDVIPLDKHRNPVNMKRRTPFRTLL 672

Qy        780 GSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLV 839
Db        673 TKFPELASFVMDNCIE-----KSKE----ETDSTQSVA-YNFEFLDDTYMMRCVSDDGTG 722

Qy        840 E--MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPY 877
Db        723 EQLIGCK-------SAYDEDFKLEKDAQSYASNYDRVYKY 755


RESULT 21
D82654
ankyrin-like protein XF1640 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82654
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82654
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1058 <SIM>
A;Cross-references: GB:AE003991; GB:AE003849; NID:g9106696; PIDN:AAF84449.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1640

  Query Match             7.8%;  Score 451;  DB 2;  Length 1058;
  Best Local Similarity   26.0%;  Pred. No. 5.1e-20;
  Matches  242;  Conservative   97;  Mismatches  348;  Indels  242;  Gaps   36;

Qy          3 GRKDVVEYLLQNGASV------QARDDGGLIPL----------------------HNAC 33
Db        204 GRVDRALQLLEAGADPHALPFDNARDQRSLMVLAAILPDLRLLRELIVRRVDVNQRHRGM 263

Qy         34 S----------FGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAE 83
Db        264 TALLAATRDSWHGRSEAVQTLLANGADPRVVDADRNTPLHHAARSSDPGVAALLRDAAAE 323

Qy         84 PTIRNTDGRTALDLADPSAKAVLTGEYKKDE-LLESARSGNEEKMMALLTPLNVNCHASD 142
Db        324 LDALNCDG------LSPLAAACQVGNWRLAKFLLE--RGAQPE-------PIG------- 361

Qy        143 GRKSTPLHLAAG---YNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVK 199
```

```
Db        362 ---GTPVLLAAAGTDEDDPTGVLLLLKHKARVDARDRLRRTALHEAAEAGHSEIVKALLE 418

Qy        200 HGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKER 259
Db        419 AAANLEVRDMAGRTPLLEAARQARIVVLECLLSHKADRLAVDSEGRNAVFLA---CLSER 475

Qy        260 LAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH---PQTHETALHC----AAASP-YP 311
Db        476 VT-------------LPLIRRLLEIGVPVVADKHGHRPVDIAAKVGCWSIVSLLDPDYP 521

Qy        312 KRKQICE------------LLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEA 358
Db        522 LPVVVSESVSVGGVLDRPPLTLLRDG------------LQLGSFGSLSELIELC----- 563

Qy        359 KVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG----FTALQMGNENVQQL 414
Db        564 ---SAEELG-TLLHDPHFALNPDVLDWLLTHGAAPTVPNSSGNVPMFSLLARGIEAVPAL 619

Qy        415 ---LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYN 471
Db        620 KVFLRHGVS------------PAGAGGLGLWLAACMQCDAASRSLE-------QFAC--- 657

Qy        472 RVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLH 531
Db        658 ------ELLEHGADPLMPSPDGDSPLVLTVRFGWLRLQQALLEAGVNCEVRDSYGMTALH 711

Qy        532 EAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLL-----------RG 579
Db        712 QATVLGREAALKLLVMHGANPDARTPDGQTSLGMALSSGRRDLATWLDWRSWPLPRRPLR 771

Qy        580 DAALLDAAKKG---CLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHG 636
Db        772 QADVPDAAMRGDTEAVRRLLDLGLP--VDAVDAQG--CTALLHAAGGGHLQVVRLLLMRG 827

Qy        637 ADVNAQDKGGLIPL----------------------HN----------AASYGHVDVAA 663
Db        828 ADLQHAANNGATPLSAAVSMRQNEIIPALLDAGAPLEHRFPEEMTVLMLAAALGFPDIVM 887

Qy        664 LLIKYNACVNATDKWAFTPLHEAAQKG-----RTQLCAL---LLAHGADPTLKNQEGQTP 715
Db        888 RLIAAGADVHASNVQGFAALHCAALYGFSARDKTRFLALLDALLLAGANPGQLADGKITP 947

Qy        716 LDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVR---SPGATADALSSGPSSPSSLSAA 772
Db        948 LLL--------LLGARAERGAV--CDEQVVLAGVERLLEEGDTLDVVDTRGFGPLHLAAL 997

Qy        773 SSLDNL------SGSFSELSSVVSSSGTE 795
Db        998 HGLPLLVKCLLRAGADSERRDALNRSPRE 1026


RESULT 22
AE2149
hypothetical protein all2748 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE2149
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-426 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74447.1; PID:g17131841; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2748
```

```
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-031-485-35

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 9.3e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db        178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 237

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db        238 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 297

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db        298 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 357

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db        358 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 417

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db        418 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 477

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db        478 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 537

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db        538 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 596

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db        597 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db        657 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db        717 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 776

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db        777 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 836

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db        837 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 896

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db        897 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 956

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db        957 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db        1017 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 1061

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db        1062 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1121

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db        1122 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1181

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db        1182 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db        1242 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1301

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db        1302 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1361

Qy       2582 A 2582
              |
Db       1362 A 1362


RESULT 24
US-09-031-485-36/c
; Sequence 36, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
```

```
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
US-09-841-835-12

  Query Match             3.2%;  Score 121.2;  DB 4;  Length 523;
  Best Local Similarity   74.3%;  Pred. No. 1.1e-21;
  Matches  153;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
             || ||  | || |||||| ||||| | | | | || || ||| | || ||| | || |||
Db          1 GGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACC 60

Qy       2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
             || |||||| ||||||| ||| | ||  | ||||| || |||||||  | ||||| ||
Db         61 CAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCT 120

Qy       2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
             || ||| |  || | || |||||||||||| |||||||| ||||| || ||  | ||||||
Db        121 GAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCT 180

Qy       2370 TTACATAATGCAGCATCTTACGGGCA 2395
              | |||||||| |||||||| ||| |
Db        181 CTTCATAATGCGGCATCTTATGGGTA 206


RESULT 22
US-09-220-132-30
; Sequence 30, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  METHODS AND COMPOSITONS FOR THE IDENTIFICATION AND ASSESSMENT
;  TITLE OF INVENTION:  OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCE
;  FILE REFERENCE:  07334-074001
;  CURRENT APPLICATION NUMBER: US/09/220,132
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  US 60/079,303
;  PRIOR FILING DATE:  1998-03-25
;  PRIOR APPLICATION NUMBER:  US 60/068,821
;  PRIOR FILING DATE:  1997-12-24
;  NUMBER OF SEQ ID NOS:  191
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 14770
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-220-132-30

  Query Match             2.6%;  Score 97.4;  DB 4;  Length 14770;
  Best Local Similarity   44.8%;  Pred. No. 1.2e-14;
  Matches  569;  Conservative    0;  Mismatches  646;  Indels   54;  Gaps    3;

Qy       1363 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1422
             ||||  |   ||| |||||||      ||  ||   |||   | ||| || ||  |
Db       1133 AGCAGGTGGTAGAAATGTTGCTTGATCGAGCTGCCCCCATTCTTTCAAAAACCAAGAATG 1192

Qy       1423 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1482
               ||  ||||  ||||| |||| | |   | ||||   |    || |  |  |
Db       1193 GATTATCTCCATTGCACATGGCCACACAAGGGGATCATTTAAACTGCGTCCAGCTTCTCC 1252

Qy       1483 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1542
             |  | ||| | | |   ||  ||| | |    |||       | |||| | ||||||   |
Db       1253 TCCAGCATAATGTACCCGTGGATGATGTCACCAATGACTACCTGACTGCCCTACACGTGG 1312

Qy       1543 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1602
             ||||  | ||||| |||    ||      |   | ||| ||     |     | | || |
Db       1313 CTGCCCACTGTGGCCATTACAAAGTTGCCAAGGTTCTCTTGGATAAGAAAGCTAACCCCA 1372

Qy       1603 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGA----------------A 1644
              |    | | ||||  | ||||||||  || | || || |                |
Db       1373 ATGCCAAAGCCCTGAATGGCTTTACCCCTCTTCATATCGCCTGCAAGAAGAATCGAATTA 1432

Qy       1645 ATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACA 1704
              | | ||  | ||  |  ||    || | |   |  | ||| |  | |||| | ||
Db       1433 AAGTAATGGAACTCCTTCTGAAACACGGTGCATCCATCCAAGCTGTAACCGAGTCGGGCC 1492

Qy       1705 GACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTG 1764
                 | | | |||           |   | ||| | |   ||     |
Db       1493 TTACCCCAATCCATGTTGCTGCCTTCATGGGGCATGTAAATATTGTATCACAACTAATGC 1552

Qy       1765 TTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAG 1824
              |||  | |          ||  | | | |       ||| |||| ||  | ||||
Db       1553 ATCATGGAGCCTCACCAAACACCACCAATGTGAGAGGAGAAACAGCACTGCACATGGCAG 1612

Qy       1825 CTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATG 1884
              || | |   |  ||      || |||   |||||  ||||  | ||||||  | ||  | |
Db       1613 CTCGCTCCGGCCAAGCTGAAGTTGTGCGGTATCTGGTACAAGACGGAGCTCAGGTAGAAG 1672

Qy       1885 CTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAG 1944
              |||||  |||  |  | ||   |||  | ||||  | ||     ||  |   ||
Db       1673 CTAAAGCTAAGGATGACCAAACACCACTCCACATTTCAGCCCGACTGGGGAAAGCAGACA 1732

Qy       1945 TTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTA 2004
              | | |||  ||  |  | || || |||          |||| |||          |  | |
Db       1733 TAGTACAACAGCTGTTGCAGCAAGGGGCATCTCCAAATGCAGCCACAACTTCTGGGTACA 1792

Qy       2005 CACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGC 2064
              | ||  | ||      | ||  | | ||  | | ||  |     |  | |  | | |
Db       1793 CCCCACTTCACCTTTCCGCCCGAGAGGGGCATGAGGATGTGGCCGCGTTCCTTTTGGATC 1852

Qy       2065 ATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCT---TGTTA 2121
              |||| ||          | | || || | | | ||| |||||||  |  || |   | |
Db       1853 ATGGAGCGTCTTTATCTATAACAACAAAGAAAGGATTTACTCCTCTTCATGTGGCAGCAA 1912

Qy       2122 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 2181
              || |||| |    ||  |    |||| ||         |   ||  |  | |||||||
Db       1913 AATATGGAAAGCTTGAAGTCGCCAATCTCCTGCTACAGAAAAGTGCATCTCCAGATGCTG 1972

Qy       2182 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAA----- 2236
              |  |||| |  |  || |  | || |        |   |  ||||||   ||
Db       1973 CTGGGAAGAGCGGGCTAACACCACTGCATGTAGCTGCACATTACGATAATCAGAAAGTGG 2032

Qy       2237 ----------------------------TTGCCGCGATACCCAAGGCAGACATTCAACAC 2268
                                          | ||  ||  |   |    ||  || | |
Db       2033 CCCTTCTGCTTTTGGACCAAGGAGCCTCACCTCACGCAGCCGCAAAGAATGGTTATACGC 2092

Qy       2269 CTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACG 2328
              |  | ||  | || ||      | ||  |  | ||  | ||      | ||  ||  | |
Db       2093 CACTGCACATCGCTGCCAAAAAGAACCAGATGGACATAGCGACAACTCTGCTGGAATATG 2152

Qy       2329 GAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTT 2388
              | |||||||  || ||          |||| ||  |  |  | |||   |||||
Db       2153 GTGCTGATGCCAACGCAGTTACCCGGCAAGGAATTGCTTCCGTCCATCTCGCAGCTCAGG 2212

Qy       2389 ACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGG 2448
              | ||||| || ||  |  | || ||  |     |||||   ||| ||  |
Db       2213 AAGGGCACGTGGACATGGTGTCGCTGCTCCTCGGTAGAAATGCGAATGTGAACCTGAGCA 2272

Qy       2449 ACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTG 2508
              | ||  | |  | || ||  | ||    || || ||| | |  |||   | |   |
Db       2273 ATAAGAGCGGCCTGACCCCACTCCATTTGGCTGCTCAAGAAGATCGAGTGAATGTGGCAG 2332

Qy       2509 CTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTT 2568
              |  |  || ||| || |||  |  |  |  ||  ||| | |||||
Db       2333 AAGTCCTCGTAAACCAAGGGGCTCATGTGGACGCCCAGACAAAGATGGGATACACACCAC 2392

Qy       2569 TAGATTTAG 2577
              | || | |
Db       2393 TGCATGTGG 2401


RESULT 23
US-09-031-485-35
; Sequence 35, Application US/09031485
```

```
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-9

  Query Match             10.0%;  Score 380.4;  DB 4;  Length 384;
  Best Local Similarity   99.7%;  Pred. No. 3.1e-89;
  Matches  381;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1701 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 62

Qy       1761 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 122

Qy       1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 182

Qy       1881 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 242

Qy       1941 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 2000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 302

Qy       2001 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 362

Qy       2061 CAGCATGGTGCAGACCCTACCA 2082
             ||||||||||||||||||||| |
Db        363 CAGCATGGTGCAGACCCTACAA 384


RESULT 20
US-09-196-387-12
; Sequence 12, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  523 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-196-387-12

  Query Match             3.2%;  Score 121.2;  DB 3;  Length 523;
  Best Local Similarity   74.3%;  Pred. No. 1.1e-21;
  Matches  153;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
             || ||  | || |||||| ||||| | | | | || || ||| | || ||| | || |||
Db          1 GGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACC 60

Qy       2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
             || |||||| ||||||| ||| | ||  | ||||| || |||||||  | ||||| ||
Db         61 CAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCT 120

Qy       2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
             || ||| |  || | || |||||||||| |||||||| ||||| || ||  | ||||||
Db        121 GAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCT 180

Qy       2370 TTACATAATGCAGCATCTTACGGGCA 2395
              | |||||||| |||||||| ||| |
Db        181 CTTCATAATGCGGCATCTTATGGGTA 206


RESULT 21
US-09-841-835-12
; Sequence 12, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  523 base pairs
;      TYPE:  nucleic acid
```

```
Db         40 AAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAAAA 1


RESULT 17
US-09-833-381-841
; Sequence 841, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 841
;   LENGTH: 1069
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1069)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-841

  Query Match           12.3%;  Score 467.6;  DB 4;  Length 1069;
  Best Local Similarity   98.8%;  Pred. No. 1.1e-111;
  Matches  492;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1662 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1721
             | | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CCCTTCCTAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 626

Qy       1722 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 686

Qy       1782 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 746

Qy       1842 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 806

Qy       1902 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1961
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 866

Qy       1962 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2021
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 926

Qy       2022 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2081
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACA 986

Qy       2082 AAAAAAAACA-GGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATAT 2140
             |||||||||| |||||||||||||||||||||||||||||||||||||| |||||||
Db        987 AAAAAAAACAGGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGAT-CAGATAT 1045

Qy       2141 TCAAGATCTGCTTAGGGG 2158
             ||||||||||||||||||
Db       1046 TCAAGATCTGCTTAGGGG 1063


RESULT 18
US-09-833-381-1153
; Sequence 1153, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1153
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1153

  Query Match           12.1%;  Score 461;  DB 4;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-110;
  Matches  461;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        572 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 64

Qy        632 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 124

Qy        692 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 184

Qy        752 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 244

Qy        812 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 304

Qy        872 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 364

Qy        932 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 424

Qy        992 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 1032
              |||||||||||||||||||||||||||||||||||||||||
Db        425 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 465


RESULT 19
US-09-972-115A-9
; Sequence 9, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 384
```

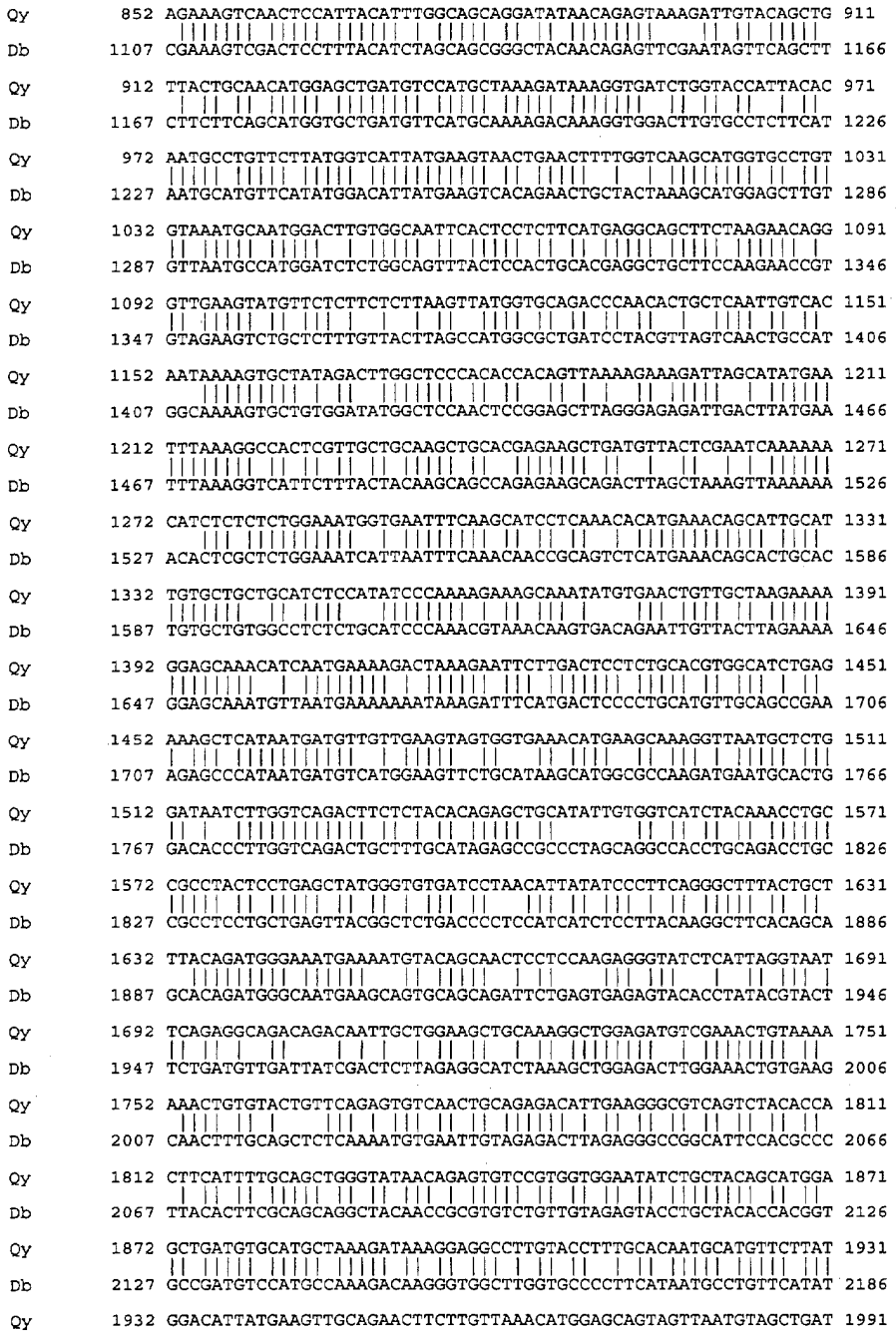
```
Db       3267 AGGAAGGACGCAGCTGTGCGCCCTCCTCCTAGCGCATGGTGCAGACCCCACCATGAAGAA 3326
Qy       2656 ------------------------------------------------------TGCTCAA 2662
Db       3327 CCAGGAAGGCCAGACGCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGAT 3386
Qy       2663 TGGTGTGAGAAGCCCAGGAGCCACTGC---------------------------------- 2689
Db       3387 AGATGCCATGCCCCCAGAGGCCTTACCTACCTGTTTTAAACCTCAGGCTACTGTAGTGAG 3446
Qy       2690 AGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGA 2749
Db       3447 TGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGCAGCATAGA 3506
Qy       2750 CAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGG 2809
Db       3507 CAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCAGGGGATGG 3566
Qy       2810 TGCTTCCAGTTTGGAGAAAAAGGAG------GTTCCAGGAGTAGATTTTAGCATAACTCA 2863
Db       3567 CGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAATATCAGCCA 3626
Qy       2864 ATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCAC 2923
Db       3627 ATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAACAGATTAC 3686
Qy       2924 TTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2983
Db       3687 ACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATCAATGCATA 3746
Qy       2984 TGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCT 3043
Db       3747 TGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAAGGCAC 3806
Qy       3044 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 3103
Db       3807 CAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTTGCTCC 3866
Qy       3104 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 3163
Db       3867 AGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAG 3926
Qy       3164 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 3223
Db       3927 AGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGT 3986
Qy       3224 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 3283
Db       3987 TGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAA 4046
Qy       3284 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3343
Db       4047 TCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTAT 4106
Qy       3344 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 3403
Db       4107 TCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTA 4166
Qy       3404 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 3463
Db       4167 TTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGG 4226
Qy       3464 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 3523
Db       4227 CTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTCTGTAG 4286
Qy       3524 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 3583
Db       4287 AGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCC 4346
Qy       3584 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 3643
Db       4347 AGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGT 4406
```

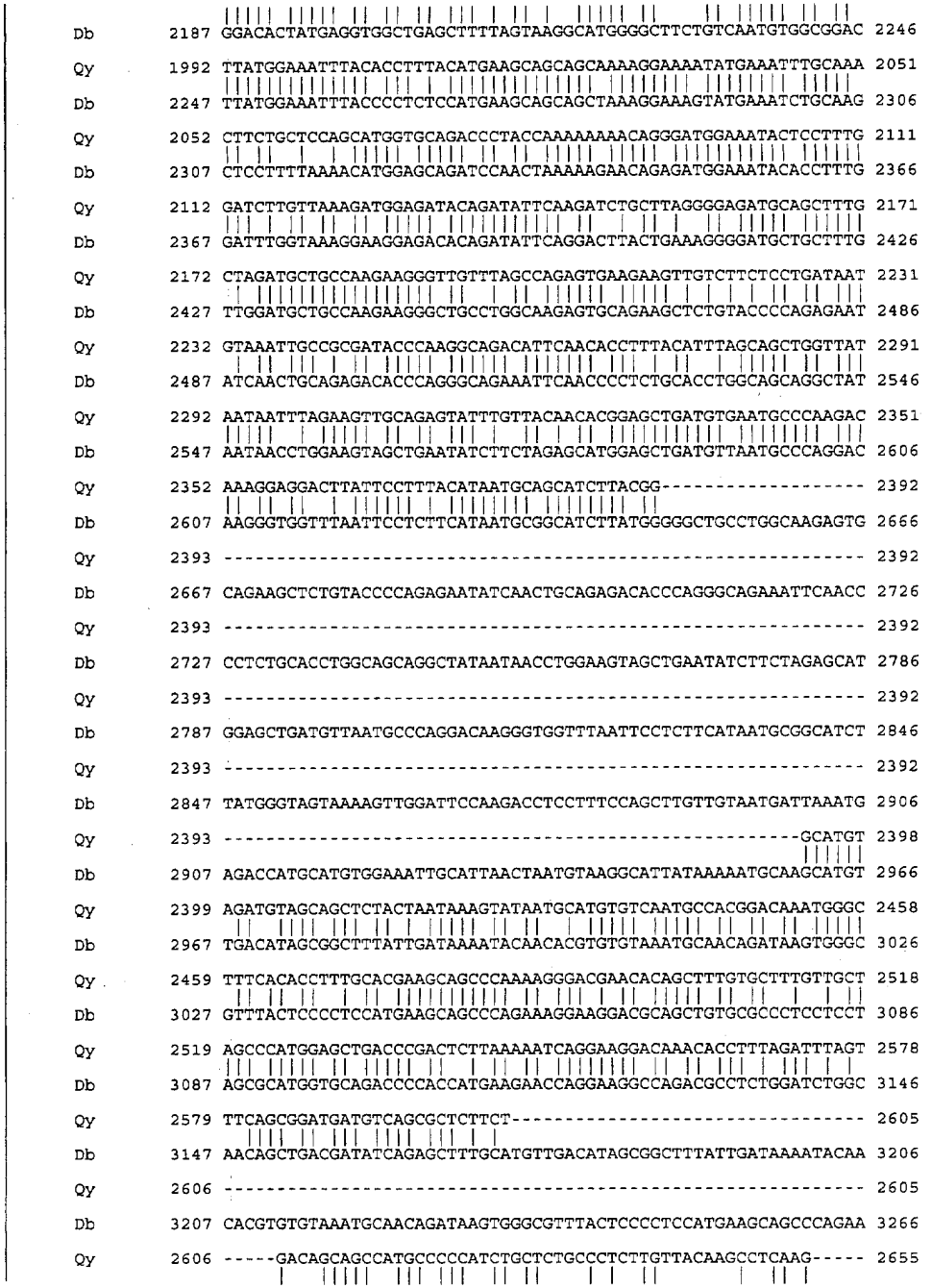
```
Qy       3644 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 3703
Db       4407 CATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATGAAGCC 4466
Qy       3704 TGAAG 3708
Db       4467 AGAAG 4471
```

RESULT 16
US-09-972-115A-10/c
; Sequence 10, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-115A-10

Query Match 13.6%; Score 518.4; DB 4; Length 520;
Best Local Similarity 99.8%; Pred. No. 4.4e-125;
Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        677 CATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGT 736
Db        520 CATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGT 461
Qy        737 GCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGA 796
Db        460 GCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGA 401
Qy        797 AAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAA 856
Db        400 AAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAA 341
Qy        857 GTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACT 916
Db        340 GTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACT 281
Qy        917 GCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGC 976
Db        280 GCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGC 221
Qy        977 CTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAA 1036
Db        220 CTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAA 161
Qy       1037 TGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGA 1096
Db        160 TGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGA 101
Qy       1097 AGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAA 1156
Db        100 AGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAA 41
Qy       1157 AAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAA 1196
```





```
Db       3747 TGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAAGGCAC 3806

Qy       3044 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 3103

Db       3807 CAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTTGCTCC 3866

Qy       3104 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 3163

Db       3867 AGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAG 3926

Qy       3164 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 3223

Db       3927 AGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGT 3986

Qy       3224 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 3283

Db       3987 TGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAA 4046

Qy       3284 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3343

Db       4047 TCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTAT 4106

Qy       3344 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 3403

Db       4107 TCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTA 4166

Qy       3404 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 3463

Db       4167 TTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGG 4226

Qy       3464 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 3523

Db       4227 CTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTCTGTAG 4286

Qy       3524 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 3583

Db       4287 AGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCC 4346

Qy       3584 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 3643

Db       4347 AGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGT 4406

Qy       3644 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 3703

Db       4407 CATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATGAAGCC 4466

Qy       3704 TGAAG 3708

Db       4467 AGAAG 4471


RESULT 15
US-09-841-835-9
; Sequence 9, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4657 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2855
US-09-841-835-9

  Query Match           35.9%;  Score 1369.4;  DB 4;  Length 4657;
  Best Local Similarity  63.1%;  Pred. No. 0;
  Matches 2501;  Conservative    0;  Mismatches  956;  Indels  508;  Gaps    5;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311

Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371

Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431

Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491

Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551

Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611

Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671

Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731

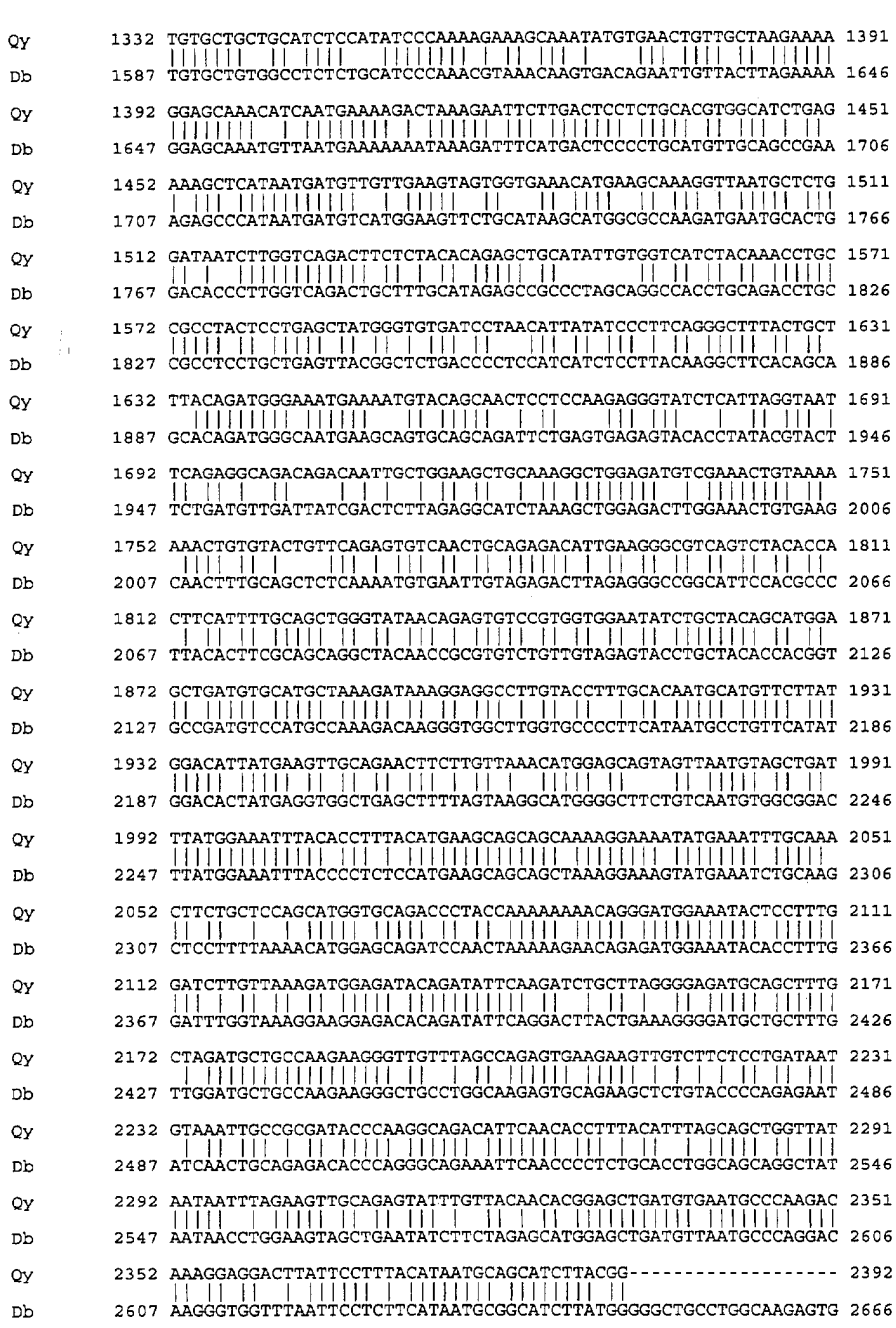
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791

Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851

Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106
```


Qy 1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
Db 1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646
Qy 1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db 1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706
Qy 1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db 1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766
Qy 1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db 1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826
Qy 1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db 1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886
Qy 1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db 1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGAGAGTACACCTATACGTACT 1946
Qy 1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAA 1751
Db 1947 TCTGATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAG 2006
Qy 1752 AAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCA 1811
Db 2007 CAACTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCC 2066
Qy 1812 CTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGA 1871
Db 2067 TTACACTTCGCAGCAGGCTACAACCGCGTGTCTGTTGTAGAGTACCTGCTACACCACGGT 2126
Qy 1872 GCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTAT 1931
Db 2127 GCCGATGTCCATGCCAAAGACAAGGGTGGCTTGGTGCCCCTTCATAATGCCTGTTCATAT 2186
Qy 1932 GGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1991
Db 2187 GGACACTATGAGGTGGCTGAGCTTTTAGTAAGGCATGGGGCTTCTGTCAATGTGGCGGAC 2246
Qy 1992 TTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAA 2051
Db 2247 TTATGGAAATTTACCCCTCTCCATGAAGCAGCAGCTAAAGGAAAGTATGAAATCTGCAAG 2306
Qy 2052 CTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTG 2111
Db 2307 CTCCTTTTAAAACATGGAGCAGATCCAACTAAAAAGAACAGAGATGGAAATACACCTTTG 2366
Qy 2112 GATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTG 2171
Db 2367 GATTTGGTAAAGGAAGGAGACACAGATATTCAGGACTTACTGAAAGGGGATGCTGCTTTG 2426
Qy 2172 CTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAAT 2231
Db 2427 TTGGATGCTGCCAAGAAGGGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAAT 2486
Qy 2232 GTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTAT 2291
Db 2487 ATCAACTGCAGAGACACCCAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTAT 2546
Qy 2292 AATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGAC 2351
Db 2547 AATAACCTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGAC 2606
Qy 2352 AAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGG---------- 2392
Db 2607 AAGGGTGGTTTAATTCCTCTTCATAATGCGGCATCTTATGGGGGCTGCCTGGCAAGAGTG 2666
Qy 2393 ---------- 2392
Db 2667 CAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAATTCAACC 2726
Qy 2393 ---------- 2392
Db 2727 CCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTTCTAGAGCAT 2786
Qy 2393 ---------- 2392
Db 2787 GGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCTCTTCATAATGCGGCATCT 2846
Qy 2393 ---------- 2392
Db 2847 TATGGGTAGTAAAAGTTGGATTCCAAGACCTCCTTTCCAGCTTGTTGTAATGATTAAATG 2906
Qy 2393 ----------GCATGT 2398
Db 2907 AGACCATGCATGTGGAAATTGCATTAACTAATGTAAGGCATTATAAAAATGCAAGCATGT 2966
Qy 2399 AGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGC 2458
Db 2967 TGACATAGCGGCTTTATTGATAAAATACAACACGTGTGTAAATGCAACAGATAAGTGGGC 3026
Qy 2459 TTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCT 2518
Db 3027 GTTTACTCCCCTCCATGAAGCAGCCCAGAAAGGAAGGACGCAGCTGTGCGCCCTCCTCCT 3086
Qy 2519 AGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGT 2578
Db 3087 AGCGCATGGTGCAGACCCCACCATGAAGAACCAGGAAGGCCAGACGCCTCTGGATCTGGC 3146
Qy 2579 TTCAGCGGATGATGTCAGCGCTCTTCT---------- 2605
Db 3147 AACAGCTGACGATATCAGAGCTTTGCATGTTGACATAGCGGCTTTATTGATAAAATACAA 3206
Qy 2606 ---------- 2605
Db 3207 CACGTGTGTAAATGCAACAGATAAGTGGGCGTTTACTCCCCTCCATGAAGCAGCCCAGAA 3266
Qy 2606 -----GACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAG----- 2655
Db 3267 AGGAAGGACGCAGCTGTGCGCCCTCCTCCTAGCGCATGGTGCAGACCCCACCATGAAGAA 3326
Qy 2656 ----------TGCTCAA 2662
Db 3327 CCAGGAAGGCCAGACGCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGAT 3386
Qy 2663 TGGTGTGAGAAGCCCAGGAGCCACTGC---------- 2689
Db 3387 AGATGCCATGCCCCCAGAGGCCTTACCTACCTGTTTTAAACCTCAGGCTACTGTAGTGAG 3446
Qy 2690 AGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGA 2749
Db 3447 TGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGCAGCATAGA 3506
Qy 2750 CAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGG 2809
Db 3507 CAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCAGGGGATGG 3566
Qy 2810 TGCTTCCAGTTTGGAGAAAAAGGAG------GTTCCAGGAGTAGATTTTAGCATAACTCA 2863
Db 3567 CGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAATATCAGCCA 3626
Qy 2864 ATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCAC 2923
Db 3627 ATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAACAGATTAC 3686
Qy 2924 TTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2983
Db 3687 ACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATCAATGCATA 3746
Qy 2984 TGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCT 3043

```
Qy       3429 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3488
Db       4026 AACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCA 4085

Qy       3489 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3548
Db       4086 TGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTG 4145

Qy       3549 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3608
Db       4146 CAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGA 4205

Qy       3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668
Db       4206 CCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATAC 4265

Qy       3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
Db       4266 CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4305
```

RESULT 14
US-09-196-387-9
; Sequence 9, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4657 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..2855
US-09-196-387-9

Query Match 35.9%; Score 1369.4; DB 3; Length 4657;
Best Local Similarity 63.1%; Pred. No. 0;

Matches 2501; Conservative 0; Mismatches 956; Indels 508; Gaps 5;

```
Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

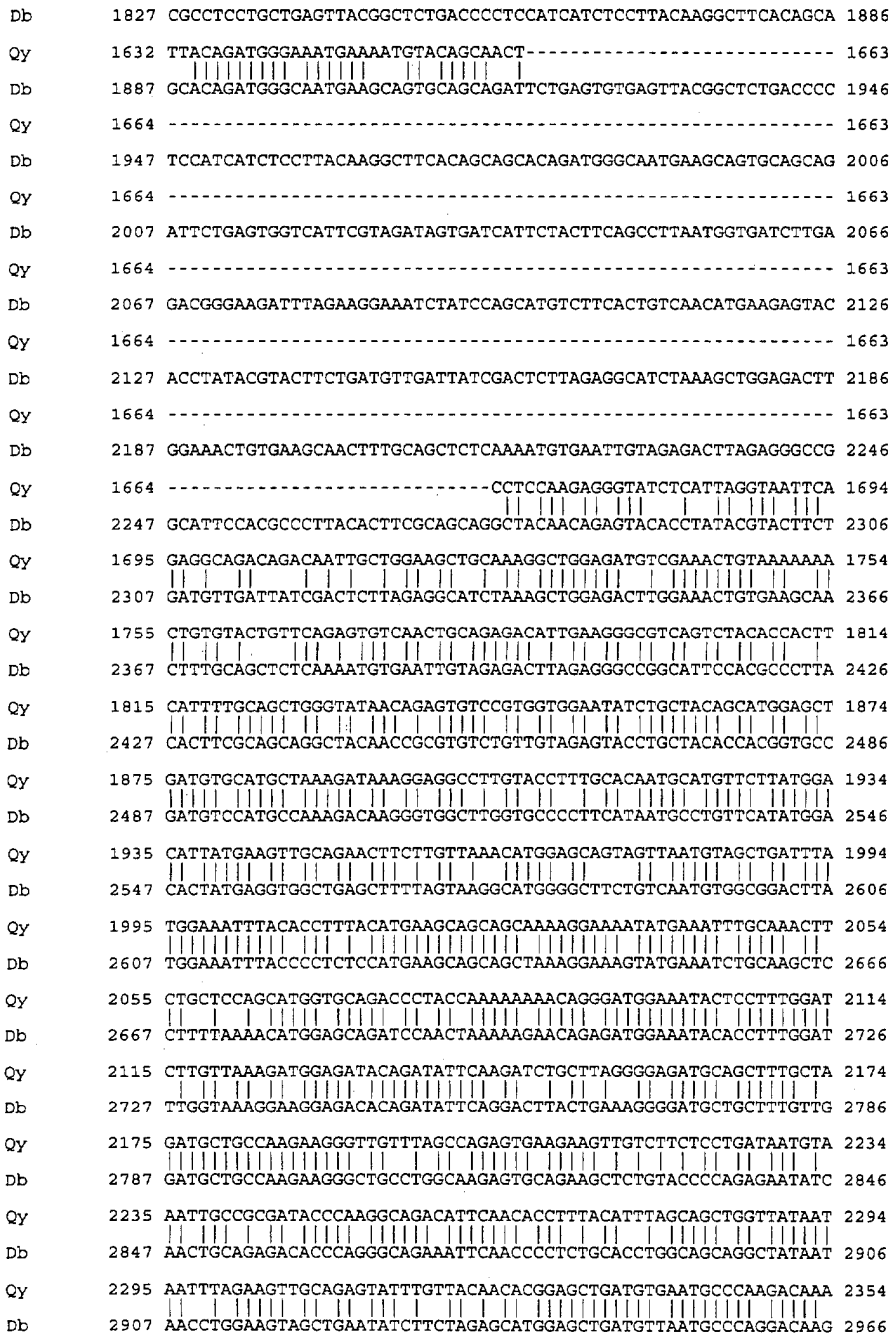
Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

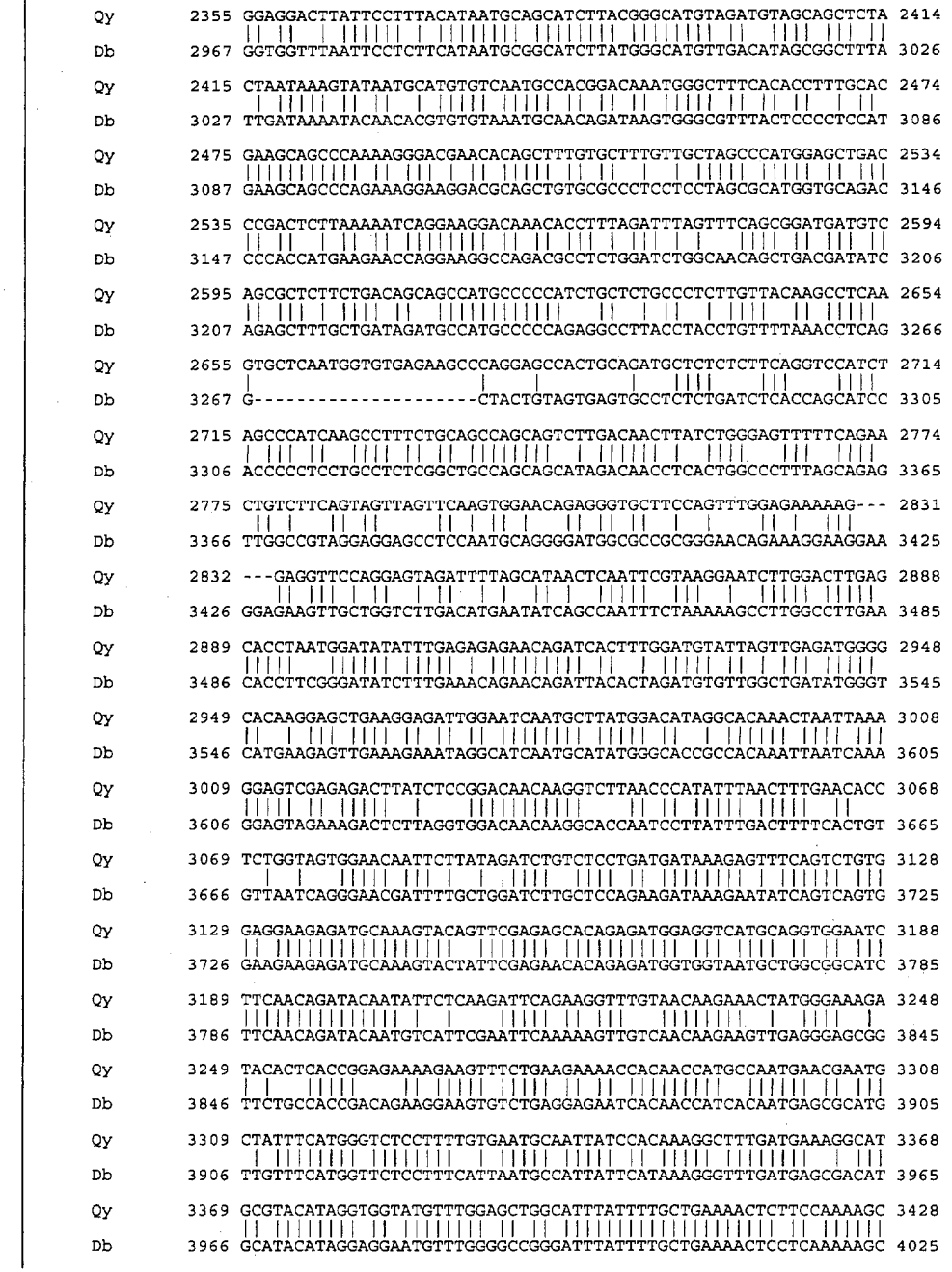
Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586
```





```
RESULT 13
US-09-841-835-7
; Sequence 7, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4491 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2027
US-09-841-835-7

  Query Match             39.9%;  Score 1521.2;  DB 4;  Length 4491;
  Best Local Similarity   65.4%;  Pred. No. 0;
  Matches 2498;  Conservative    0;  Mismatches  938;  Indels  384;  Gaps    3;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | | || |   || |||     ||||  || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              ||||| || ||||| |||||||||    || || ||||||||||||     | |  || ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||       |||| ||| ||| |||| ||  | || ||||||||||| ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
               |||| || || || ||| | || || |||| ||||| |||||||| || |||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
```

```
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
               | ||| | || || || || || ||||| | ||||| ||||| ||||| ||||| |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||| |||||||| |||| |||| |||||||||||| | ||||| |||||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              |||| ||| || || || ||||| |   |||| |||| | |||||||| || || |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              || || ||||| |||||||| |||||||| |||||| |||||   ||| ||||||||| |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              ||||||||| | |||||| |||| || ||  ||||||| || ||||| ||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              ||||||| ||||| |||||| | ||||| || || ||||||||     || || |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
               | || || ||||| |||||||| ||||| ||||| |||||||  || || ||  | ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              ||||| ||||| ||||| |||||||||| || ||||  |  | |||||||| || |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              || ||||| |||||  | |||||  || ||||| || || ||||| ||||| ||||||  |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              || ||||| || ||| |  |  | || |||| || || || ||  |  |||| || ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
                 |||||||| | ||  |||||| || ||  || | | || ||||  | ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              |||||||  || || || |||||  |||||||  ||   |  ||  | | ||||||
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
                 |||  |||||||||  | |||||||  || || ||  |||||||||||| ||||
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

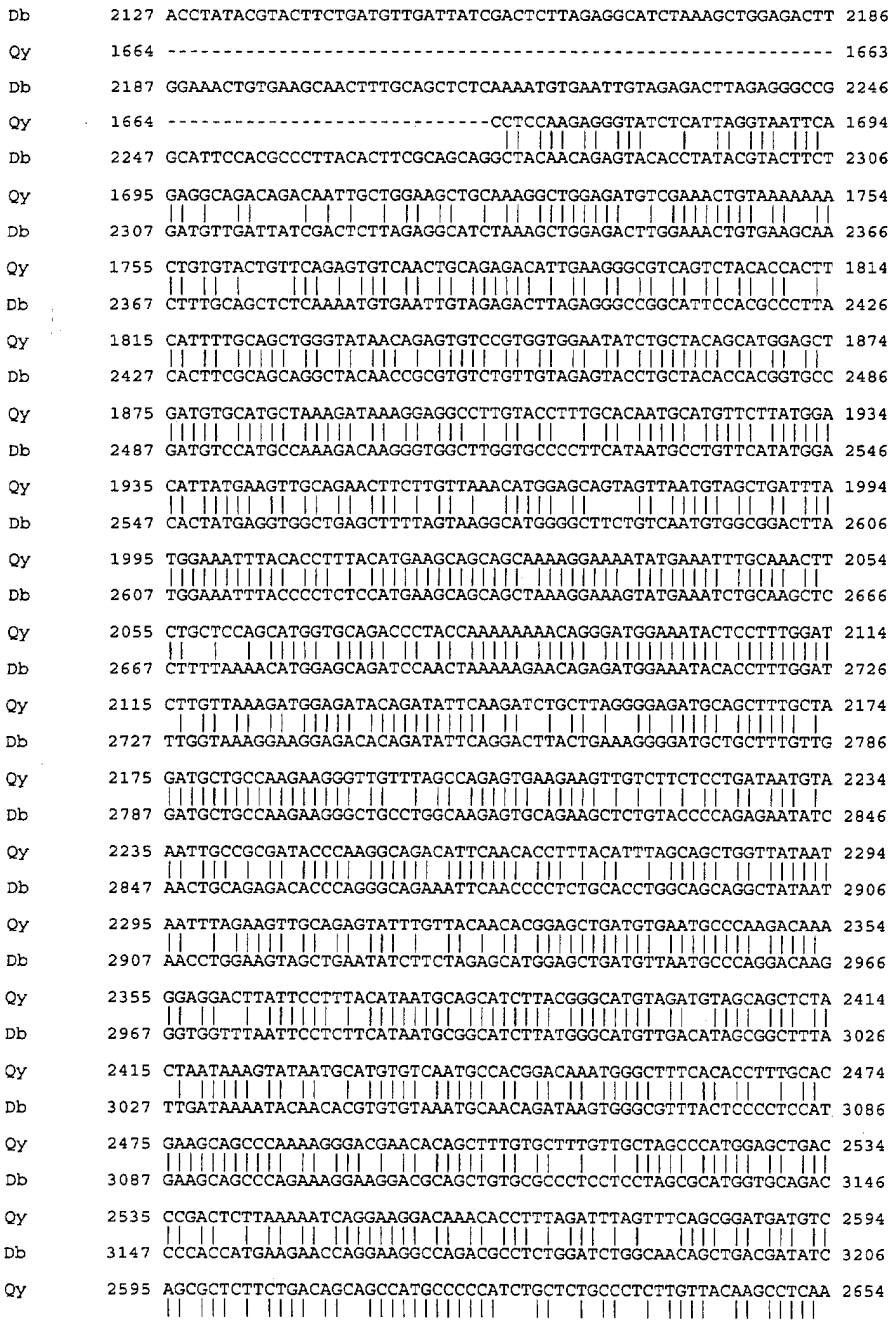
Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
              |||||||  || ||||   |||||||| | || |||  |   ||| |||| || ||||||
Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
              |||||||  | |||||||| | |||||| ||| |||||| ||||| || ||| | |||
Db       1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
              | ||| ||||||||||  | |||||  ||   ||  |||| || ||| | ||||| |||
Db       1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
              || | |||||||||||| || | || ||||| ||          || || || ||||||
Db       1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
              ||||| || ||||| || || | ||| ||   ||| || ||| | || ||||| || ||
```


Db 2127 ACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTT 2186
Qy 1664 ---------- 1663
Db 2187 GGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCG 2246
Qy 1664 ----------CCTCCAAGAGGGTATCTCATTAGGTAATTCA 1694
Db 2247 GCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACAGAGTACACCTATACGTACTTCT 2306
Qy 1695 GAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAA 1754
Db 2307 GATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAGCAA 2366
Qy 1755 CTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTT 1814
Db 2367 CTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCCTTA 2426
Qy 1815 CATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCT 1874
Db 2427 CACTTCGCAGCAGGCTACAACCGCGTGTCTGTTGTAGAGTACCTGCTACACCACGGTGCC 2486
Qy 1875 GATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGA 1934
Db 2487 GATGTCCATGCCAAAGACAAGGGTGGCTTGGTGCCCCTTCATAATGCCTGTTCATATGGA 2546
Qy 1935 CATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTA 1994
Db 2547 CACTATGAGGTGGCTGAGCTTTTAGTAAGGCATGGGGCTTCTGTCAATGTGGCGGACTTA 2606
Qy 1995 TGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTT 2054
Db 2607 TGGAAATTTACCCCTCTCCATGAAGCAGCAGCTAAAGGAAAGTATGAAATCTGCAAGCTC 2666
Qy 2055 CTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGAT 2114
Db 2667 CTTTTAAAACATGGAGCAGATCCAACTAAAAAGAACAGAGATGGAAATACACCTTTGGAT 2726
Qy 2115 CTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTA 2174
Db 2727 TTGGTAAAGGAAGGAGACACAGATATTCAGGACTTACTGAAAGGGGATGCTGCTTTGTTG 2786
Qy 2175 GATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTA 2234
Db 2787 GATGCTGCCAAGAAGGGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATC 2846
Qy 2235 AATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAAT 2294
Db 2847 AACTGCAGAGACACCCAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAAT 2906
Qy 2295 AATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAA 2354
Db 2907 AACCTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGACAAG 2966
Qy 2355 GGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTA 2414
Db 2967 GGTGGTTTAATTCCTCTTCATAATGCGGCATCTTATGGGCATGTTGACATAGCGGCTTTA 3026
Qy 2415 CTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCAC 2474
Db 3027 TTGATAAAATACAACACGTGTGTAAATGCAACAGATAAGTGGGCGTTTACTCCCCTCCAT 3086
Qy 2475 GAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGAC 2534
Db 3087 GAAGCAGCCCAGAAAGGAAGGACGCAGCTGTGCGCCCTCCTCCTAGCGCATGGTGCAGAC 3146
Qy 2535 CCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTC 2594
Db 3147 CCCACCATGAAGAACCAGGAAGGCCAGACGCCTCTGGATCTGGCAACAGCTGACGATATC 3206
Qy 2595 AGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAA 2654
Db 3207 AGAGCTTTGCTGATAGATGCCATGCCCCCAGAGGCCTTACCTACCTGTTTTAAACCTCAG 3266
Qy 2655 GTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCT 2714
Db 3267 G----------CTACTGTAGTGAGTGCCTCTCTGATCTCACCAGCATCC 3305
Qy 2715 AGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAA 2774
Db 3306 ACCCCCTCCTGCCTCTCGGCTGCCAGCAGCATAGACAACCTCACTGGCCCTTTAGCAGAG 3365
Qy 2775 CTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG--- 2831
Db 3366 TTGGCCGTAGGAGGAGCCTCCAATGCAGGGGATGGCGCCGCGGGAACAGAAAGGAAGGAA 3425
Qy 2832 ---GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAG 2888
Db 3426 GGAGAAGTTGCTGGTCTTGACATGAATATCAGCCAATTTCTAAAAAGCCTTGGCCTTGAA 3485
Qy 2889 CACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGG 2948
Db 3486 CACCTTCGGGATATCTTTGAAACAGAACAGATTACACTAGATGTGTTGGCTGATATGGGT 3545
Qy 2949 CACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAA 3008
Db 3546 CATGAAGAGTTGAAAGAAATAGGCATCAATGCATATGGGCACCGCCACAAATTAATCAAA 3605
Qy 3009 GGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACC 3068
Db 3606 GGAGTAGAAAGACTCTTAGGTGGACAACAAGGCACCAATCCTTATTTGACTTTTCACTGT 3665
Qy 3069 TCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTG 3128
Db 3666 GTTAATCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCAGTG 3725
Qy 3129 GAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATC 3188
Db 3726 GAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGCATC 3785
Qy 3189 TTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGA 3248
Db 3786 TTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGG 3845
Qy 3249 TACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATG 3308
Db 3846 TTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGCATG 3905
Qy 3309 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 3368
Db 3906 TTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGACAT 3965
Qy 3369 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 3428
Db 3966 GCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAAAGC 4025
Qy 3429 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3488
Db 4026 AACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCA 4085
Qy 3489 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3548
Db 4086 TGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTG 4145
Qy 3549 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3608
Db 4146 CAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGA 4205
Qy 3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668
Db 4206 CCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATAC 4265
Qy 3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
Db 4266 CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4305

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4491 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2027
US-09-196-387-7

  Query Match             39.9%;  Score 1521.2;  DB 3;  Length 4491;
  Best Local Similarity   65.4%;  Pred. No. 0;
  Matches 2498;  Conservative    0;  Mismatches  938;  Indels  384;  Gaps    3;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | | ||  |  || |||     ||||   || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              |||||  ||  |||||  |||||||||    ||  ||  ||||||||||||     |  ||  ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||        |||| ||| ||| |||| ||  |  || ||||||||||| ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              ||||  ||  ||  ||  |||  |  ||  ||  ||||  |||||  |||||||| ||  |||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              ||  ||  ||  ||  ||  ||||||||||  ||  ||||||||  ||||||||  ||  ||  |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              |  |||  |  ||  ||  ||  ||  |||||  |  |||||  |||||  |||||  |||||  |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||  ||||||||  |||||  |||||  ||||||||||||  |  |||||  ||||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              ||||  |||  ||  ||  ||  |||||  |    ||||  ||||  |  ||||||||  ||  ||  |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              ||  ||  |||||  ||||||||  ||||||||  ||||||  |||||  |||  ||||||||  |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              |||||||||  |  ||||||  ||||  ||  ||  |||||||  ||  |||||  ||||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              |||||||  |||||  ||||||  |||||  ||  ||  ||||||||     ||  ||  |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
              |  ||  ||  |||||  ||||||||  |||||  |||||  |||||||   ||  ||  |  ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              |||||  |||||  |||||  ||||||||||  ||  |||||  |  |  ||||||||  ||  |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              ||  |||||  |||||  |  |||||  ||  |||||  ||  ||  |||||  |||||  ||||||  |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              ||  |||||  ||  |||  |  |  |  ||  ||||  ||  ||  ||  ||  |  ||||  ||  ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
                 |||||||||  |  ||||||  ||  ||  ||  |  |  |  ||  |||||  |  ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              |||||||  ||  ||  ||  ||  |||||  ||  |||||||  ||  |  ||  |  |  ||||||
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
                 |||  ||||||||||  |  ||||||||  ||  ||  ||  |  ||||||||||  ||||
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
              |||||||  ||  ||||  ||||||||  |  ||  |||  |  |||  ||||  ||  ||||||
Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
              |||||||  |  ||||||||  |  ||||||  |||  |||||||  |||||  ||  |||  |  ||
Db       1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
              |  |||  ||||||||||  |||||  ||  ||  ||||  |||  |  |||||  |||
Db       1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
              ||  |  ||||||||||||  ||  |  ||  |||||  ||          ||  ||  ||  ||  ||||||
Db       1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
              |||||  ||  |||||  ||  ||  |  |||  ||  ||  |||  |  ||  |||||  ||  ||
Db       1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACT---------------------------- 1663
                 |||||||||  ||||||    ||  |||||  |
Db       1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGTGAGTTACGGCTCTGACCCC 1946

Qy       1664 ------------------------------------------------------------ 1663

Db       1947 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 2006

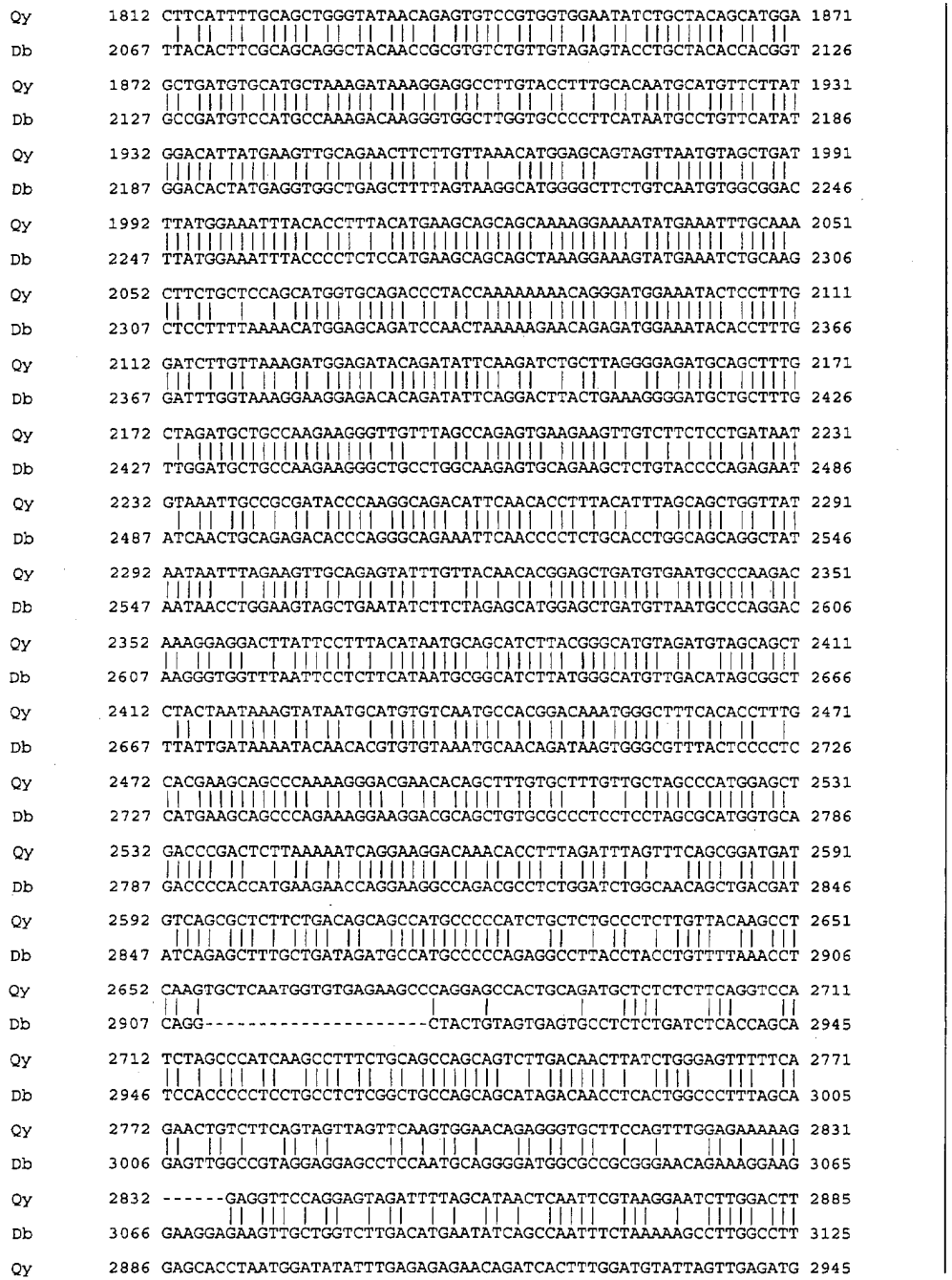
Qy       1664 ------------------------------------------------------------ 1663

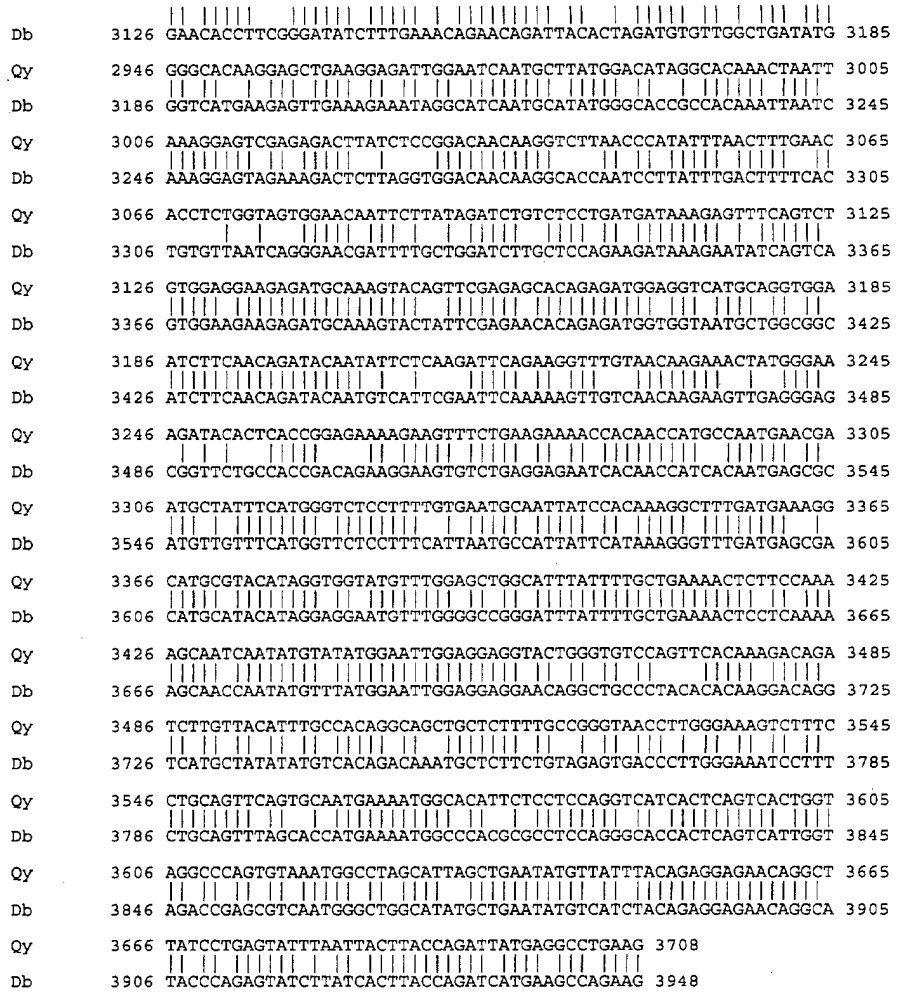
Db       2007 ATTCTGAGTGGTCATTCGTAGATAGTGATCATTCTACTTCAGCCTTAATGGTGATCTTGA 2066

Qy       1664 ------------------------------------------------------------ 1663

Db       2067 GACGGGAAGATTTAGAAGGAAATCTATCCAGCATGTCTTCACTGTCAACATGAAGAGTAC 2126

Qy       1664 ------------------------------------------------------------ 1663
```





RESULT 12
US-09-196-387-7
; Sequence 7, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

```
Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
             |||||||| ||   | |||||||||||| ||   | |||||||| || |||||||||| ||||
Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
             || || || || ||||| || |||| |||||||||| || ||||||||||||||||||||
Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
             || || |||||| | || |||||||||| |||| ||| ||||
Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948


RESULT 11
US-09-972-115A-7
; Sequence 7, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 4134
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-7

  Query Match             49.4%;  Score 1885;  DB 4;  Length 4134;
  Best Local Similarity   72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | | || |  || |||     ||||  || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              ||||| || ||||| |||||||||   || || ||||||||||||   |  ||  ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||      |||| ||| ||| |||| ||  | || ||||||||||| ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              ||||  ||  || ||  ||| |  ||  |||| ||||| |||||||| || |||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              || || || || ||||||||||| || |||||||| |||||||| || || | |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              | ||| | || || || || || ||||| | ||||| ||||| ||||| ||||| |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||| |||||||| ||||| ||||| |||||||||||| | ||||| ||||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              |||| ||| || || || ||||| |   ||||  |||| | |||||||| || || |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              || || ||||| |||||||| |||||||| ||||| |||||  ||| |||||||| |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              |||||||| | |||||| |||| || || ||||||| || ||||| ||||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              |||||| ||||| |||||| | ||||| || || ||||||||    || || |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
              | || || ||||| |||||||| ||||| ||||| |||||||  || || ||  | ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              ||||| ||||| ||||| |||||||||||| || |||||  |  | |||||||| || |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              || |||| |||||  | ||||| || ||||| || || |||||  ||||| |||||| |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              || ||||| || ||| |  |  || |||| || || || ||  | |||| || ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
              |||||||| | || ||||||| || || || | | |||| | ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              |||||||| || || || || ||||| ||  ||||||| ||  |  ||  | | ||||||
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
                 ||| ||||||||||  | |||||||| || || ||  | |||||||||||| ||||
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
              |||||||  || ||||  |||||||| | || |||  |  ||| |||| || ||||||
Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
              |||||||  | |||||||| | |||||| ||| |||||||  |||| || ||| | ||
Db       1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
              | ||| |||||||||||  | |||||  ||  || ||||  || ||| | |||||  |||
Db       1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

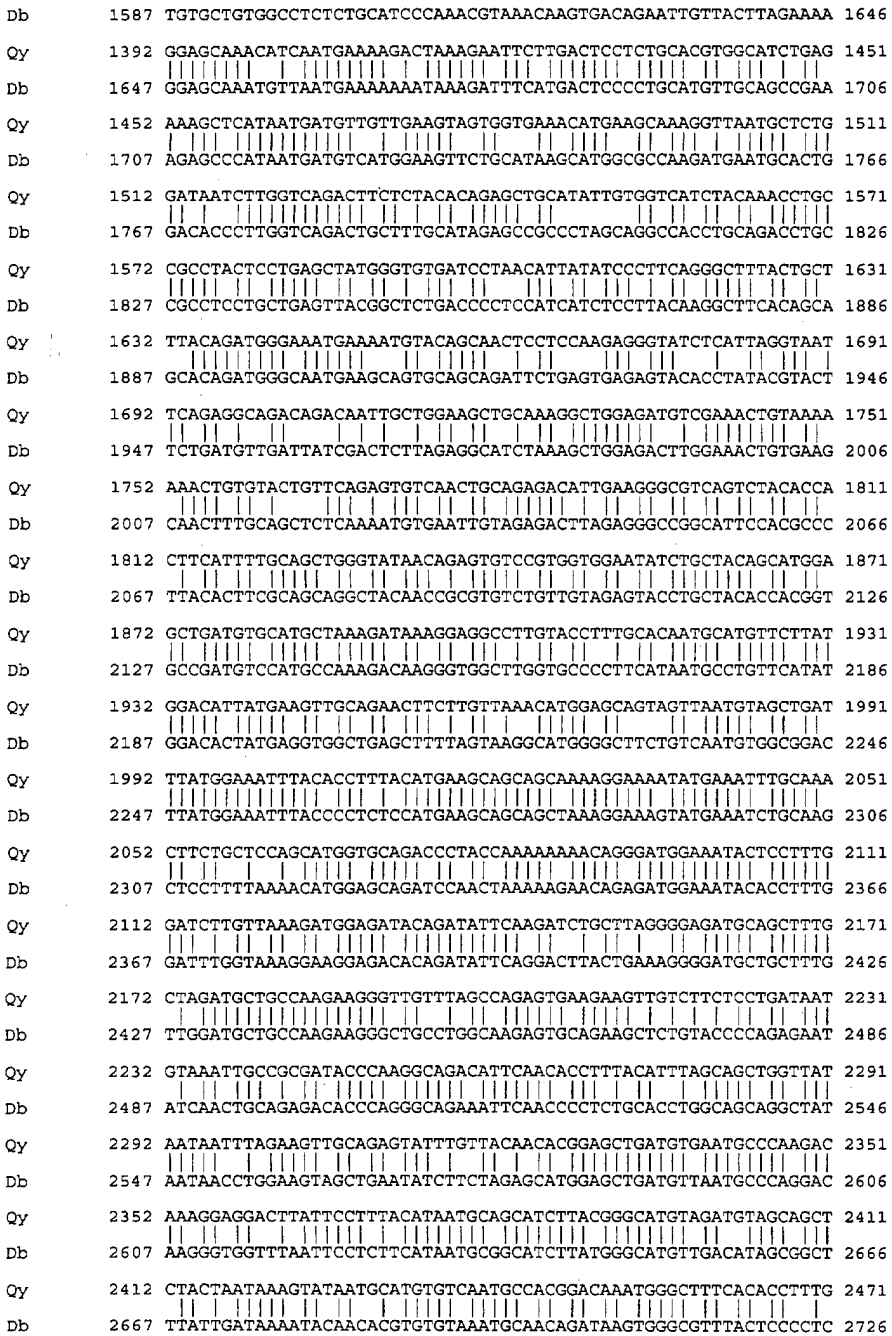
Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
              || |  |||||||||||| || | || |||||  ||      || || || || ||||||
Db       1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

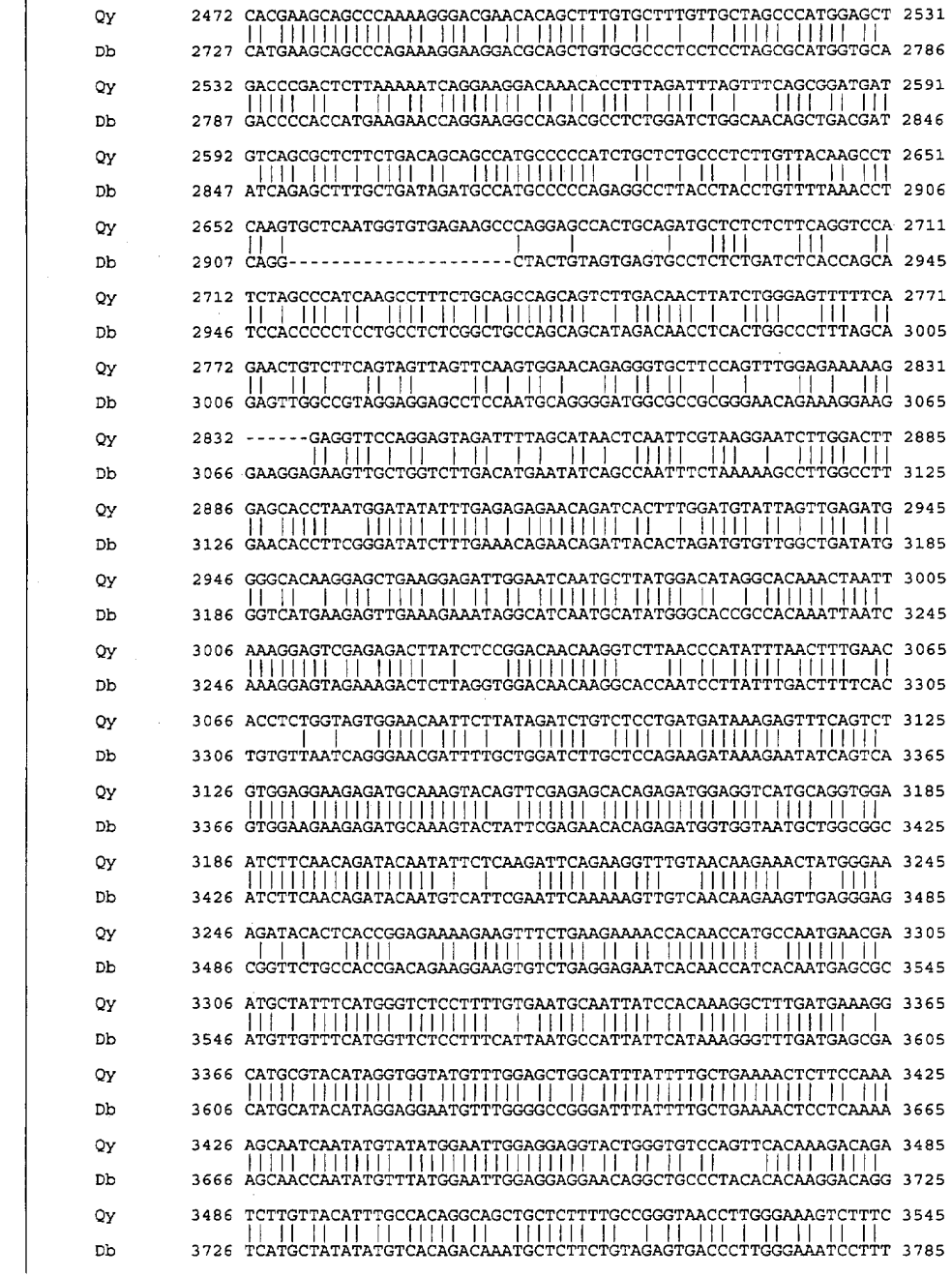
Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
              ||||| || ||||| || || | || ||   ||| || ||| | || |||||  || ||
Db       1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
              |||||||| |||||  || |||||  | ||     ||| |||    | || ||| | |
Db       1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGAGAGTACACCTATACGTACT 1946

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAA 1751
              || || |  ||  | | |  || ||  || |||||||  | |||||||| ||
Db       1947 TCTGATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAG 2006

Qy       1752 AAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCA 1811
              |||| || |   ||| | ||| || || |||||| | || || || || || ||
Db       2007 CAACTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCC 2066
```





```
Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485
Db       3666 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3725

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545
Db       3726 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3785

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948
```

```
RESULT 10
US-09-841-835-1
; Sequence 1, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4134 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-841-835-1

  Query Match             49.4%;  Score 1885;  DB 4;  Length 4134;
  Best Local Similarity   72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
```

```
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

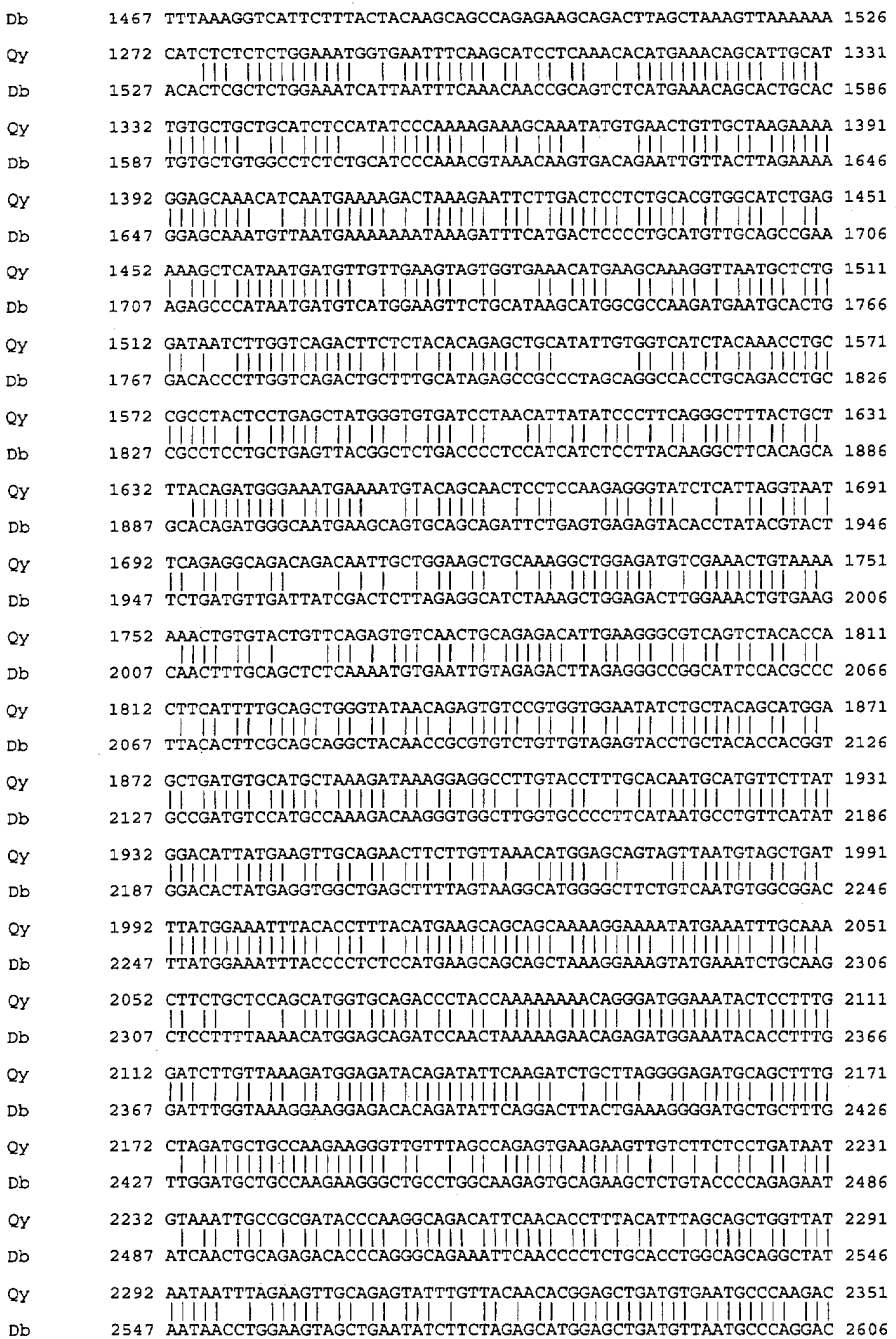
Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

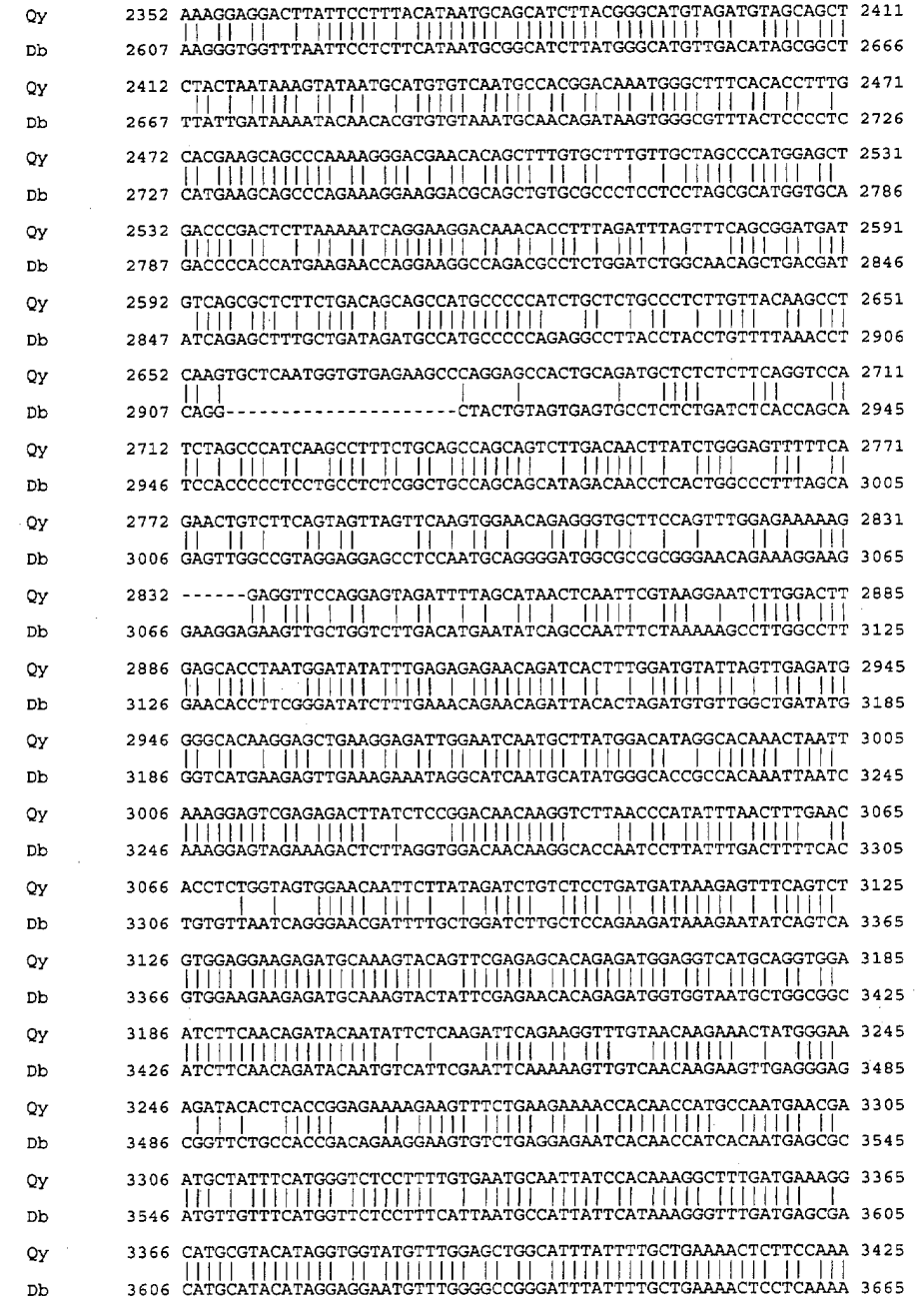
Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
```





```
Qy       3430 ATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTT 3489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3707 ATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTT 3766

Qy       3490 GTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGC 3549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3767 GTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGC 3826

Qy       3550 AGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGC 3609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3827 AGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGC 3886

Qy       3610 CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATC 3669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3887 CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATC 3946

Qy       3670 CTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAG 3729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3947 CTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAG 4006

Qy       3730 TTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACG 3789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4007 TTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACG 4066

Qy       3790 TTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||
Db       4067 TTTTACTCCTTTGCTGAAAAAAAA 4090
```

```
RESULT 9
US-09-196-387-1
; Sequence 1, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4134 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-196-387-1

  Query Match            49.4%;  Score 1885;  DB 3;  Length 4134;
  Best Local Similarity  72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | || |  || |||     ||||   || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              ||||| || ||||| |||||||||    || || ||||||||||||    | |  || ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||        |||| ||| ||| |||| ||  | || ||||||||||| ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              |||| || || || ||| | || || |||| ||||| |||||||| || ||||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              || || || || || |||||||||| || ||||||| |||||||| || || | |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              | ||| | || || || || || ||||| | ||||| ||||| ||||| ||||| |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||| |||||||| ||||| ||||| |||||||||||| | ||||| ||||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              |||| ||| || || || ||||| |  |||| |||| | |||||||| || || |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              || || ||||| |||||||| |||||||| |||||| |||| ||| ||||||||| |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              |||||||| | |||||| |||| || || |||||| || ||||| ||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              ||||||| ||||| |||||| | ||||| || || |||||||    || || |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
              | || || ||||| |||||||| ||||| ||||| |||||||  || || ||  | ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

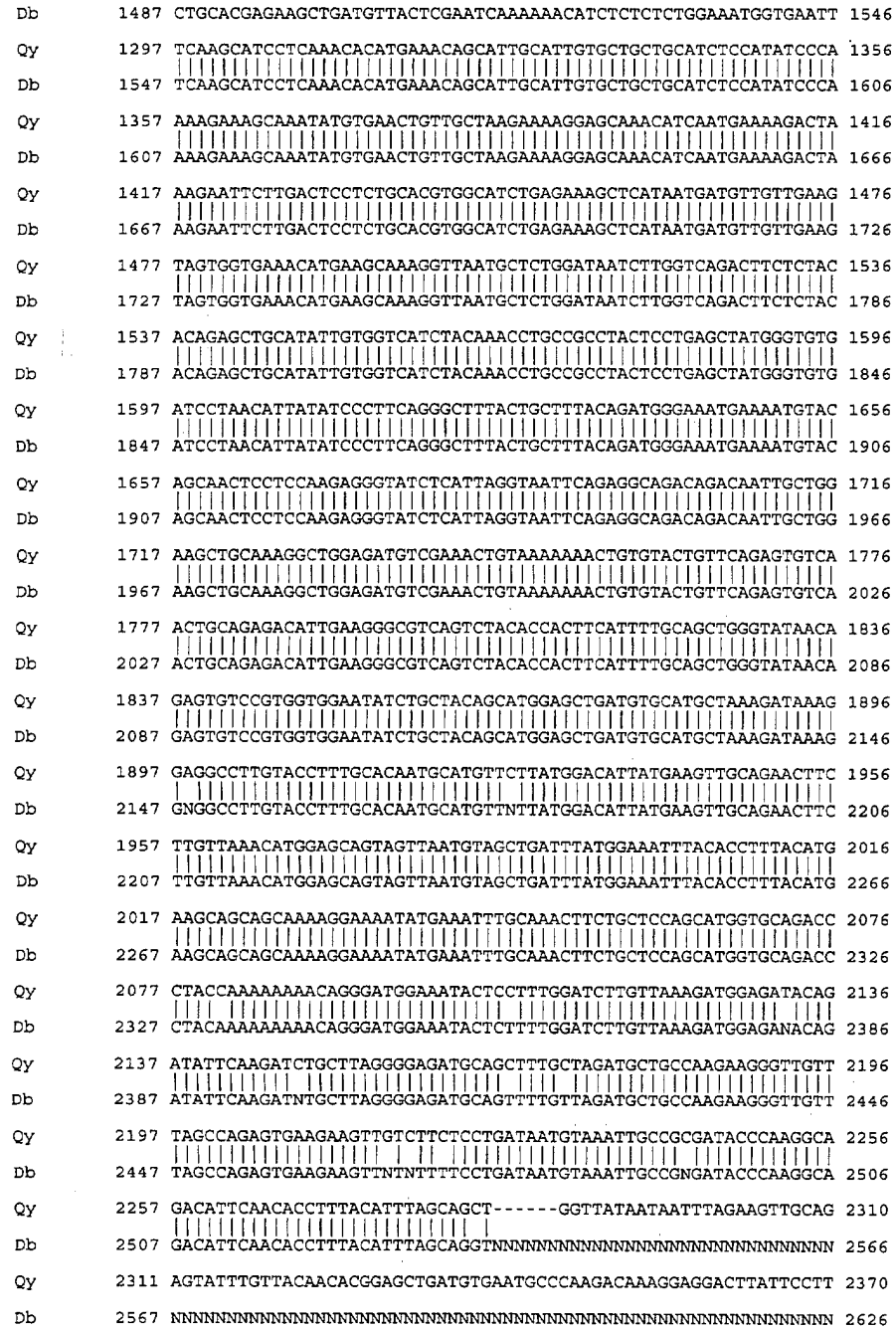
Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              ||||| ||||| ||||| |||||||||||| || |||||  |  | |||||||| || |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              || ||||| ||||  | ||||| || ||||| || || ||||| ||||| |||||| |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              || ||||| || ||| |  |  | ||  |||| || || || ||  |  |||| || ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
                |||||||| |  ||  |||||| || ||  || | |  || |||||  | ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              |||||||| || || || || ||||| ||  ||||||| ||  |  ||  | | ||||||
```


Db 1487 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1546
Qy 1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
Db 1547 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1606
Qy 1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
Db 1607 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1666
Qy 1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
Db 1667 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1726
Qy 1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
Db 1727 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1786
Qy 1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
Db 1787 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1846
Qy 1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
Db 1847 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1906
Qy 1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
Db 1907 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1966
Qy 1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
Db 1967 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 2026
Qy 1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836
Db 2027 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 2086
Qy 1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896
Db 2087 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 2146
Qy 1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956
Db 2147 GNGGCCTTGTACCTTTGCACAATGCATGTTNTTATGGACATTATGAAGTTGCAGAACTTC 2206
Qy 1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2016
Db 2207 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2266
Qy 2017 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2076
Db 2267 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2326
Qy 2077 CTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2136
Db 2327 CTACAAAAAAAAACAGGGATGGAAATACTCTTTTGGATCTTGTTAAAGATGGAGANACAG 2386
Qy 2137 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2196
Db 2387 ATATTCAAGATNTGCTTAGGGGAGATGCAGTTTTGTTAGATGCTGCCAAGAAGGGTTGTT 2446
Qy 2197 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2256
Db 2447 TAGCCAGAGTGAAGAAGTTNTNTTTTCCTGATAATGTAAATTGCCGNGATACCCAAGGCA 2506
Qy 2257 GACATTCAACACCTTTACATTTAGCAGCT------GGTTATAATAATTTAGAAGTTGCAG 2310
Db 2507 GACATTCAACACCTTTACATTTAGCAGGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2566
Qy 2311 AGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTT 2370
Db 2567 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2626
Qy 2371 TACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATG 2430
Db 2627 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2686
Qy 2431 CATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGG 2490
Db 2687 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2746
Qy 2491 GACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATC 2550
Db 2747 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2806
Qy 2551 AGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAG 2610
Db 2807 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAG 2866
Qy 2611 CAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGA 2670
Db 2867 CAGCCATGCCCCCATCTGTTCTGCCCTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGA 2926
Qy 2671 GAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTT 2730
Db 2927 GAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTT 2986
Qy 2731 CTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTA 2790
Db 2987 CTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTA 3046
Qy 2791 GTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATT 2850
Db 3047 GTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATT 3106
Qy 2851 TTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGA 2910
Db 3107 TTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTNAGA 3166
Qy 2911 GAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTG 2970
Db 3167 GAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTG 3226
Qy 2971 GAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCG 3030
Db 3227 GWATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCG 3286
Qy 3031 GACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTA 3090
Db 3287 GACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTA 3346
Qy 3091 TAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAG 3150
Db 3347 TAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAG 3406
Qy 3151 TTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCA 3210
Db 3407 TTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCA 3466
Qy 3211 AGATTCAGAAGGTTTGTAACAAGAAACTATG----------GGAAAGAT 3249
Db 3467 AGATTCAGAAGGTTTGTAACANNNNNNNNNNNGAGCCAAGATTCGGCACGAGGAAAGAT 3526
Qy 3250 ACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGC 3309
Db 3527 ACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGC 3586
Qy 3310 TATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATG 3369
Db 3587 TATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATG 3646
Qy 3370 CGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCA 3429
Db 3647 CGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCA 3706

; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (621)..(621)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (626)..(626)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (629)..(629)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (638)..(638)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (693)..(694)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (724)..(724)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (730)..(730)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (764)..(764)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (791)..(791)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2148)..(2148)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2177)..(2177)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2382)..(2382)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2466)..(2466)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2468)..(2468)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2493)..(2493)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2536)..(2860)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (3163)..(3163)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put

; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (3488)..(3499)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
US-09-972-115A-1

Query Match 79.6%; Score 3039; DB 4; Length 4493;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 3101; Conservative 1; Mismatches 375; Indels 27; Gaps 2;

```
Qy       337 AACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGA 396
            ||| | ||   |||||||||| ||||| ||||| |||| || | |||||| |||||||||
Db       587 AACCATTCCNNAGGCTGGTGACCCCTGAAAAGGTNAACANCCNCAACACGGNGGGCAGGA 646

Qy       397 AATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGC 456
            |||||||||||||||||||| |||||||||||||||||||| || |  || |||||||||
Db       647 AATCCACCCCGCTGCACTTCCCCGCAGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGC 706

Qy       457 TTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATG 516
            |||| |||||||||||| ||||| ||| | || ||||||||||||||||||||||||| ||
Db       707 TTCAAAATGGTGCAAATNTCCAANCACTTTATAATGGGGGCCTTATTCCTCTTCATANTG 766

Qy       517 CATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCA 576
            |||||||||||||||||| ||    ||||||||||||||||||||||||||||||||||||
Db       767 CATGCTCTTTTGGTCATGCTAAAANTATCAATCTCCTTTTGCGACATGGTGCAGACCCCA 826

Qy       577 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 636
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       827 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 886

Qy       637 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 696
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       887 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 946

Qy       697 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 756
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       947 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 1006

Qy       757 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 816
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1007 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 1066

Qy       817 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 876
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1067 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 1126

Qy       877 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 936
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1127 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 1186

Qy       937 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 996
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1187 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 1246

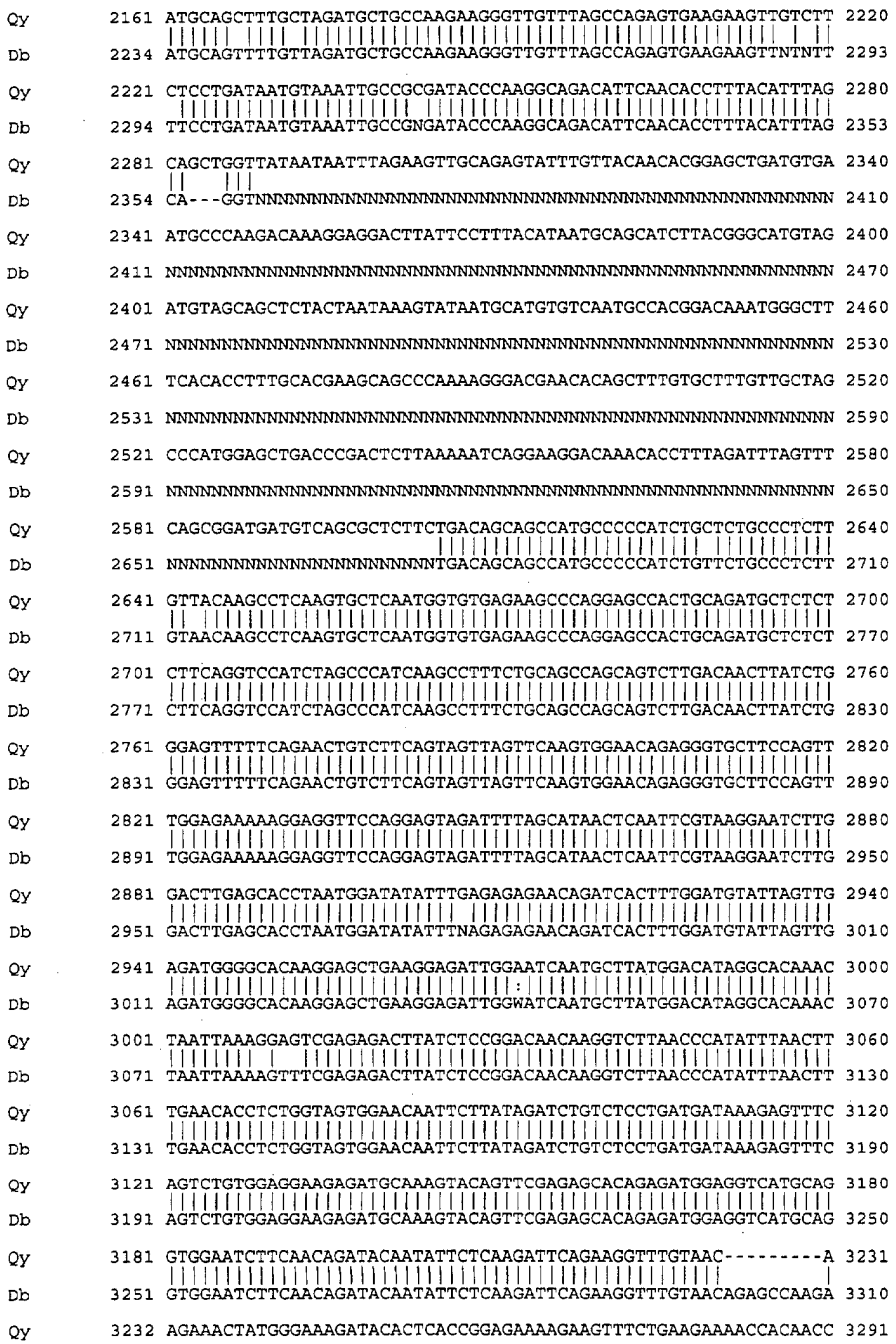
Qy       997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1247 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1306

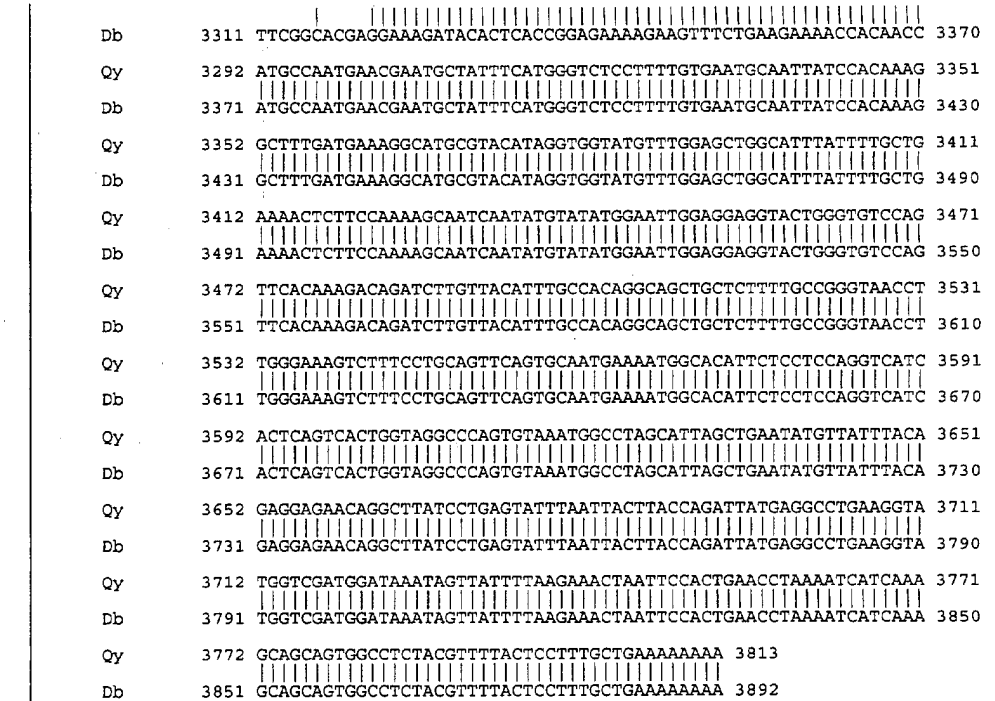
Qy      1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1307 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1366

Qy      1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1367 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1426

Qy      1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1427 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1486

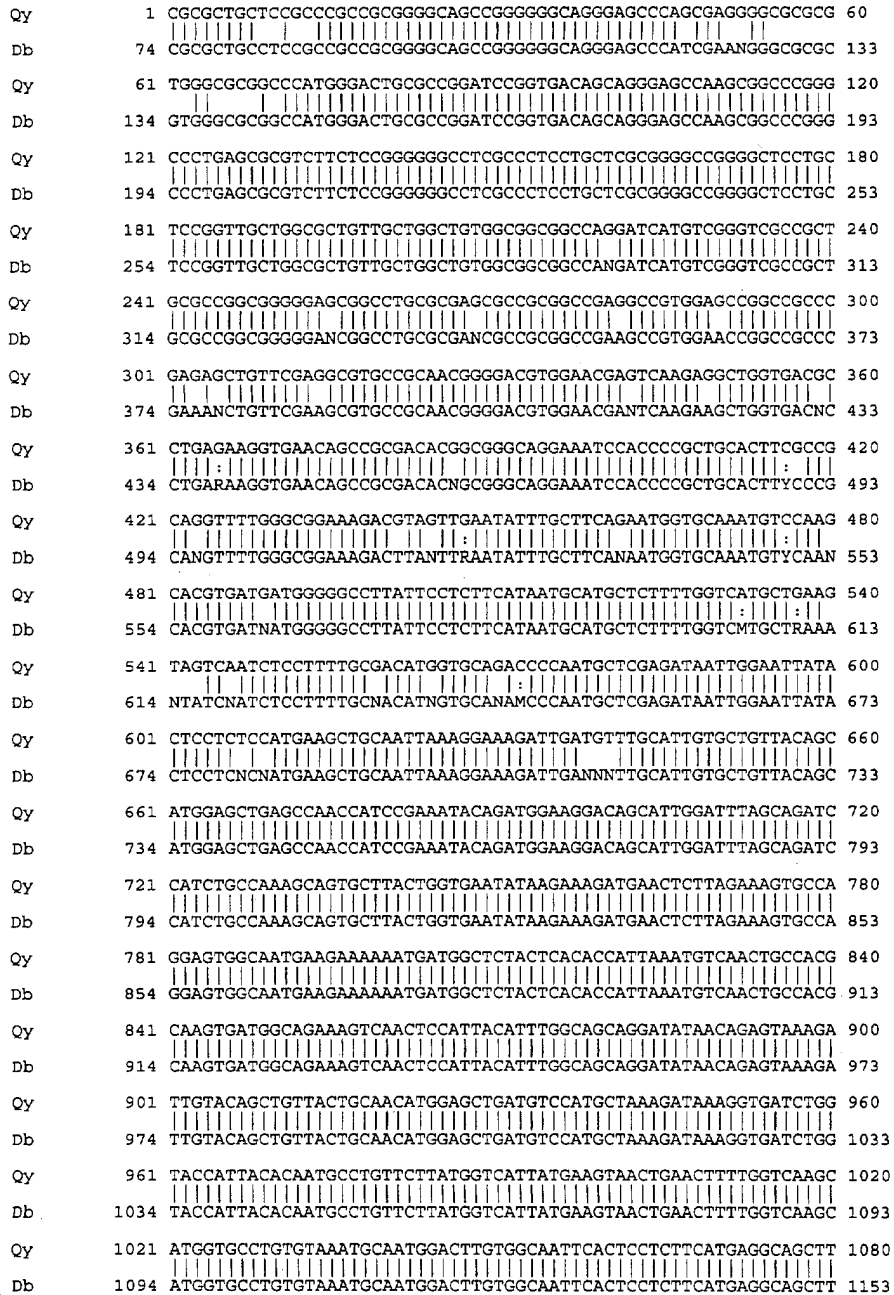
Qy      1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

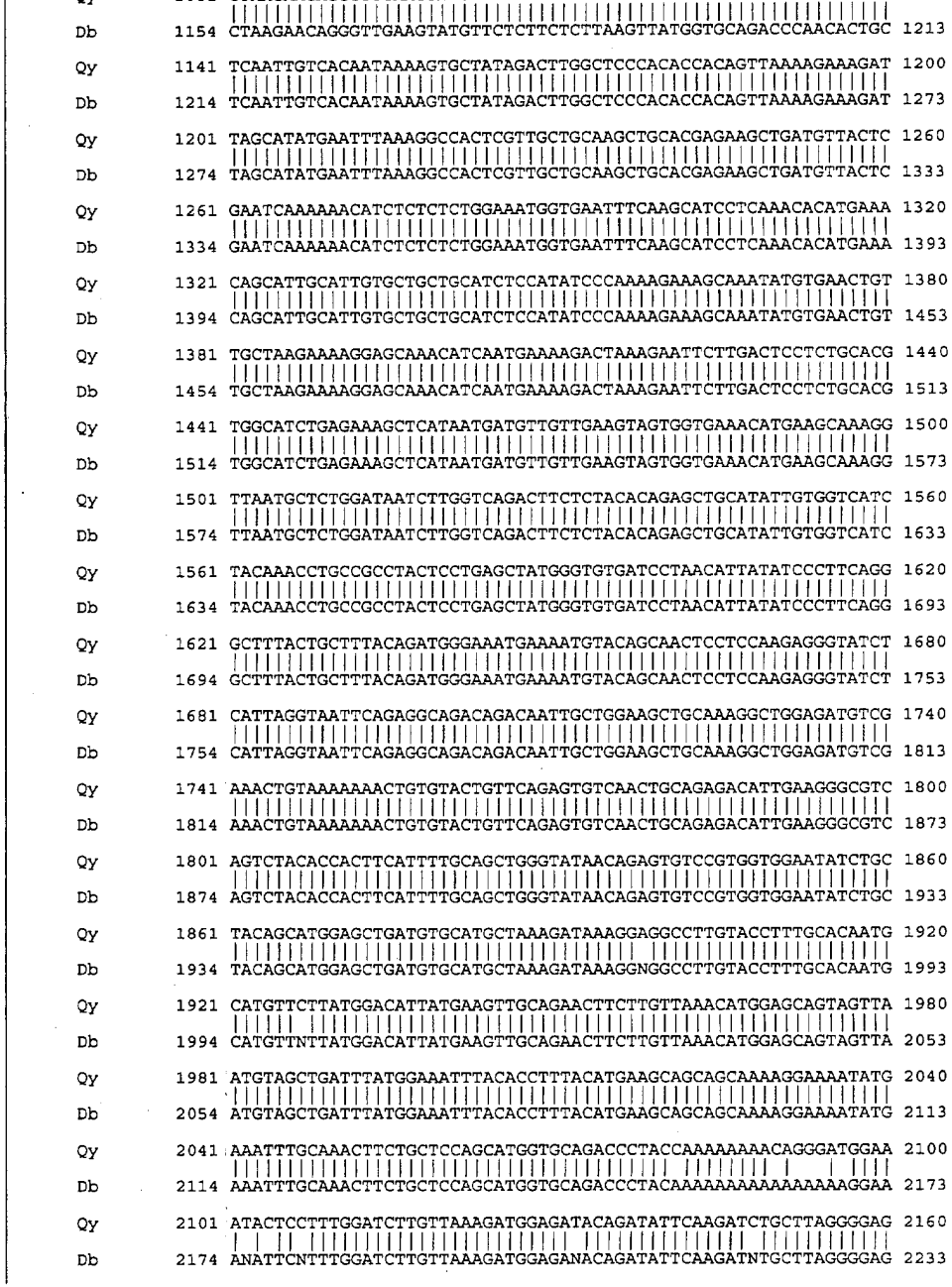




```
RESULT 8
US-09-972-115A-1
; Sequence 1, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3999)
;   OTHER INFORMATION:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(586)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (596)..(597)
```

Matches 3400; Conservative 8; Mismatches 402; Indels 12; Gaps 2;





```
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 4297
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3801)
;   OTHER INFORMATION:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (125)..(125)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (293)..(293)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (328)..(328)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (341)..(341)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (378)..(378)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (415)..(415)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (432)..(432)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (459)..(459)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (496)..(496)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (517)..(517)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (534)..(534)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (553)..(553)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (562)..(562)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (614)..(614)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (619)..(619)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (632)..(632)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (637)..(637)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (643)..(643)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (680)..(680)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (682)..(682)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (711)..(713)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (1971)..(1971)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2000)..(2000)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2175)..(2175)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2180)..(2180)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2205)..(2205)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2221)..(2221)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2289)..(2289)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2316)..(2316)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2359)..(2674)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2977)..(2977)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2291)..(2291)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
US-09-972-115A-3

  Query Match            87.6%;  Score 3341;  DB 4;  Length 4297;
  Best Local Similarity  89.0%;  Pred. No. 0;
```

```
Qy       1923 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1560

Qy       1983 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620

Qy       2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 2102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 1680

Qy       2103 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2162
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCATTATCTGCTTAGGGGAGAT 1740

Qy       2163 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 1800

Qy       2223 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 1860

Qy       2283 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 1920

Qy       2343 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 1980

Qy       2403 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2040

Qy       2463 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2100

Qy       2523 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160

Qy       2583 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2220

Qy       2643 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2280

Qy       2703 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2340

Qy       2763 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2822
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AGTTTTTCAGAACTGTCTTCATTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2400

Qy       2823 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2882
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2460

Qy       2883 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2942
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2520

Qy       2943 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 3002
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2580

Qy       3003 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 3062
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 2640

Qy       3063 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 3122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 2700

Qy       3123 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 3182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 2760

Qy       3183 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 2820

Qy       3243 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 2880

Qy       3303 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 2940

Qy       3363 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3000

Qy       3423 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3060

Qy       3483 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3120

Qy       3543 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3180

Qy       3603 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3240

Qy       3663 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3300

Qy       3723 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3360

Qy       3783 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
             |||||||||||||||||||||||||||||||||||
Db       3361 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3394
```

RESULT 7
US-09-972-115A-3
; Sequence 3, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13

```
Db       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Qy       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661

Db       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Qy       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721

Db       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Qy       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781

Db       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Qy       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816

Db       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797


RESULT 6
US-09-427-154-1
; Sequence 1, Application US/09427154
; Patent No. 6589725
; GENERAL INFORMATION:
;  APPLICANT:  Luo, Ying
;  APPLICANT:  Chan, Eva
;  APPLICANT:  Xu, Xiang
;  APPLICANT:  Huang, Betty
;  TITLE OF INVENTION:  TANKYRASE H, COMPOSTIONS INVOLVED IN THE CELL CYCLE AND
;  TITLE OF INVENTION:  METHODS OF USE
;  FILE REFERENCE:  A-68292-DJB/RMS/DAV
;  CURRENT APPLICATION NUMBER: US/09/427,154
;  CURRENT FILING DATE:  1999-10-25
;  NUMBER OF SEQ ID NOS:  9
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 3394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-427-154-1

  Query Match             88.8%;  Score 3387.6;  DB 4;  Length 3394;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3390;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        423 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 482

Db          1 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 60

Qy        483 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 542

Db         61 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 120

Qy        543 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 602

Db        121 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 180

Qy        603 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 662

Db        181 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 240

Qy        663 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 722

Db        241 GGAGCTGAGCCAACCATCCTAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 300

Qy        723 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 782

Db        301 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 360

Qy        783 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 842

Db        361 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 420

Qy        843 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 902

Db        421 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480

Qy        903 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 962

Db        481 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 540

Qy        963 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1022

Db        541 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 600

Qy       1023 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1082

Db        601 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 660

Qy       1083 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1142

Db        661 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 720

Qy       1143 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1202

Db        721 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 780

Qy       1203 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1262

Db        781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 840

Qy       1263 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1322

Db        841 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 900

Qy       1323 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1382

Db        901 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 960

Qy       1383 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1442

Db        961 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1020

Qy       1443 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1502

Db       1021 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1080

Qy       1503 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1562

Db       1081 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1140

Qy       1563 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1622

Db       1141 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1200

Qy       1623 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1682

Db       1201 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1260

Qy       1683 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1742

Db       1261 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1320

Qy       1743 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1802

Db       1321 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1380

Qy       1803 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1862

Db       1381 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1440

Qy       1863 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1922

Db       1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1500
```

```
Qy    1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Db    1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422

Qy    1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Db    1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482

Qy    1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Db    1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542

Qy    1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Db    1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602

Qy    1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Db    1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662

Qy    1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Db    1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722

Qy    1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Db    1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782

Qy    1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Db    1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842

Qy    1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Db    1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902

Qy    1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Db    1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962

Qy    1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Db    1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022

Qy    2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2101
Db    2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082

Qy    2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Db    2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142

Qy    2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db    2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202

Qy    2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db    2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262

Qy    2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db    2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322

Qy    2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db    2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382

Qy    2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db    2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442

Qy    2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db    2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502

Qy    2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db    2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562

Qy    2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Db    2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622

Qy    2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Db    2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682

Qy    2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Db    2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742

Qy    2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Db    2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802

Qy    2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Db    2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862

Qy    2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Db    2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922

Qy    2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Db    2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982

Qy    3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
Db    2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042

Qy    3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
Db    3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102

Qy    3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
Db    3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162

Qy    3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
Db    3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

Qy    3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
Db    3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

Qy    3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
Db    3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342

Qy    3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
Db    3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402

Qy    3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
Db    3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462

Qy    3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
Db    3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

Qy    3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCAGTCAGTCAC 3601
```

```
Db       2701 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2760

Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 2820

Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 2880

Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 2940

Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3000

Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3060

Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3120

Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3180

Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3240

Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3300

Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3360

Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3420

Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3480

Qy       3705 GAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||
Db       3481 GAAGGTATGGTCGATGGA 3498


RESULT 5
US-09-696-668-1
; Sequence 1, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
;  APPLICANT: Luo, Ying
;  APPLICANT:  Chan, Eva
;  APPLICANT:  Xu, Xiang
;  APPLICANT:  Huang, Betty
;  APPLICANT:  Ossovskaya, Valeria
;  TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS
;  FILE REFERENCE: A-68292-1/RMS/DHR
;  CURRENT APPLICATION NUMBER: US/09/696,668
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 09/427,154
;  PRIOR FILING DATE: 1999-10-25
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1
;   LENGTH: 3797
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-696-668-1

  Query Match           88.9%;  Score 3393.4;  DB 4;  Length 3797;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3394;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Qy        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522

Qy        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582

Qy        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642

Qy        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702

Qy        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762

Qy        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822

Qy        842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882

Qy        902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942

Qy        962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002

Qy       1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062

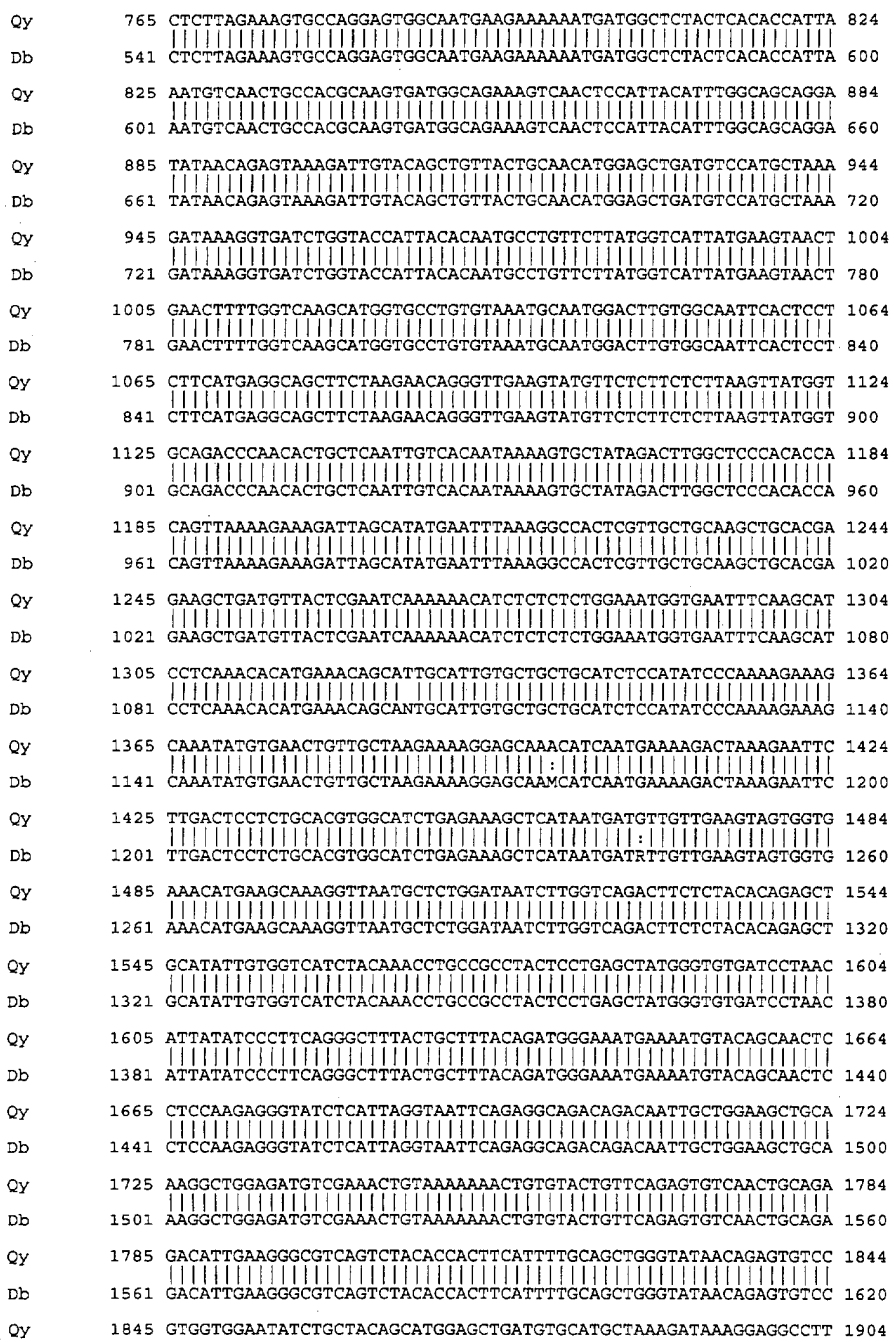
Qy       1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122

Qy       1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182

Qy       1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242

Qy       1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302

Qy       1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
```


Qy 765 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 824
Db 541 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 600
Qy 825 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
Db 601 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660
Qy 885 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 944
Db 661 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 720
Qy 945 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 1004
Db 721 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 780
Qy 1005 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 1064
Db 781 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 840
Qy 1065 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 1124
Db 841 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 900
Qy 1125 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 1184
Db 901 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 960
Qy 1185 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1244
Db 961 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1020
Qy 1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1304
Db 1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1080
Qy 1305 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364
Db 1081 CCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140
Qy 1365 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1424
Db 1141 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTC 1200
Qy 1425 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484
Db 1201 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTG 1260
Qy 1485 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544
Db 1261 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320
Qy 1545 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604
Db 1321 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380
Qy 1605 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1664
Db 1381 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1440
Qy 1665 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1724
Db 1441 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1500
Qy 1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
Db 1501 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1560
Qy 1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
Db 1561 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1620
Qy 1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
Db 1621 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTT 1680
Qy 1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
Db 1681 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1740
Qy 1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
Db 1741 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 1800
Qy 2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2084
Db 1801 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAA 1860
Qy 2085 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2144
Db 1861 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 1920
Qy 2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2204
Db 1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGA 1980
Qy 2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
Db 1981 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2040
Qy 2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
Db 2041 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2100
Qy 2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
Db 2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2160
Qy 2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
Db 2161 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220
Qy 2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
Db 2221 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2280
Qy 2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
Db 2281 TGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2340
Qy 2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
Db 2341 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2400
Qy 2625 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
Db 2401 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2460
Qy 2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
Db 2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2520
Qy 2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
Db 2521 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2580
Qy 2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
Db 2581 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640
Qy 2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
Db 2641 TTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2700
Qy 2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984

```
Qy       2790 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2588 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2647

Qy       2850 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2909
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       2648 TTTAGCATAACTCAATTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2707

Qy       2910 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2708 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2767

Qy       2970 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 3029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2768 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 2827

Qy       3030 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 3089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2828 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 2887

Qy       3090 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 3149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2888 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 2947

Qy       3150 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2948 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3007

Qy       3210 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3008 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3067

Qy       3270 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3068 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3127

Qy       3330 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3128 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3187

Qy       3390 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3188 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3247

Qy       3450 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3248 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3307

Qy       3510 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3308 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3367

Qy       3570 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3368 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3427

Qy       3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3428 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3487

Qy       3690 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3488 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3547

Qy       3750 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3548 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3607

Qy       3810 AAAA 3813
             ||||
Db       3608 AAAA 3611
```

```
RESULT 4
US-09-350-982C-4
; Sequence 4, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
;  APPLICANT:  Berthelsen, Jens
;  APPLICANT:  Toma, Salvatore
;  APPLICANT:  Isacchi, Antonella
;  TITLE OF INVENTION:  Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Rela
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  PHRM-0043
;  CURRENT APPLICATION NUMBER: US/09/350,982C
;  CURRENT FILING DATE:  1999-07-09
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 4
;  LENGTH: 3498
;  TYPE: DNA
;  ORGANISM: Artificial
;  FEATURE:
;  OTHER INFORMATION: PCR Primers
;  NAME/KEY: CDS
;  LOCATION: (1)..(3498)
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Xaa is any amino acid
;  NAME/KEY: misc_feature
;  LOCATION: (1102)..(1102)
;  OTHER INFORMATION: n is any nucleic acid
;  NAME/KEY: misc_feature
;  LOCATION: (2650)..(2650)
;  OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-4

  Query Match            91.4%;  Score 3488.8;  DB 4;  Length 3498;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 3487;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy        285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy        345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy        405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy        465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy        525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy        585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

Qy        645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy        705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 540
```

```
Db        368 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 427
Qy        630 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 689
Db        428 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 487
Qy        690 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 749
Db        488 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 547
Qy        750 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 809
Db        548 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 607
Qy        810 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 869
Db        608 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 667
Qy        870 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 929
Db        668 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 727
Qy        930 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 989
Db        728 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 787
Qy        990 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 1049
Db        788 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 847
Qy       1050 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 1109
Db        848 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 907
Qy       1110 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 1169
Db        908 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 967
Qy       1170 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229
Db        968 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027
Qy       1230 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1289
Db       1028 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1087
Qy       1290 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCA 1349
Db       1088 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCA 1147
Qy       1350 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAA 1409
Db       1148 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAA 1207
Qy       1410 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTT 1469
Db       1208 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTT 1267
Qy       1470 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1529
Db       1268 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1327
Qy       1530 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1589
Db       1328 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1387
Qy       1590 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1649
Db       1388 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1447
Qy       1650 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1709
Db       1448 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1507

Qy       1710 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1769
Db       1508 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1567
Qy       1770 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1829
Db       1568 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1627
Qy       1830 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1889
Db       1628 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1687
Qy       1890 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1949
Db       1688 GATAAAGRRGSCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1747
Qy       1950 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 2009
Db       1748 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 1807
Qy       2010 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069
Db       1808 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 1867
Qy       2070 GCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 2129
Db       1868 GCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 1927
Qy       2130 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 2189
Db       1928 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 1987
Qy       2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
Db       1988 AGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2047
Qy       2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
Db       2048 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2107
Qy       2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
Db       2108 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2167
Qy       2370 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2429
Db       2168 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2227
Qy       2430 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2489
Db       2228 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2287
Qy       2490 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2549
Db       2288 GGACGAACACAGCTTTGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2347
Qy       2550 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2609
Db       2348 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2407
Qy       2610 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2669
Db       2408 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2467
Qy       2670 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729
Db       2468 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2527
Qy       2730 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2789
Db       2528 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2587
```

```
Db       2580 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2639
Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2640 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2699
Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2700 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2759
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2760 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2819
Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       2820 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2879
Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       2880 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2939
Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       2940 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2999
Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       3000 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3059
Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3060 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3119
Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3120 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3179
Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3180 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3239
Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3240 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3299
Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3300 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3359
Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3360 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3419
Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       3420 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3479
Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       3480 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3539
Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       3540 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3599
Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       3600 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3659
Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       3660 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3719
```

```
Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       3720 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3779
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       3780 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3839
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3840 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 3
US-09-350-982C-3
; Sequence 3, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
; APPLICANT: Berthelsen, Jens
; APPLICANT: Toma, Salvatore
; APPLICANT: Isacchi, Antonella
; TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Rela
; TITLE OF INVENTION: Same
; FILE REFERENCE: PHRM-0043
; CURRENT APPLICATION NUMBER: US/09/350,982C
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 4512
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: PCR Primers
; NAME/KEY: misc_feature
; LOCATION: (1124)..(1124)
; OTHER INFORMATION: n is any nucleic acid
; NAME/KEY: misc_feature
; LOCATION: (2672)..(2672)
; OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-3

Query Match 94.2%; Score 3594.8; DB 4; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3593; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```
Qy        210 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 269
Db          8 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 67
Qy        270 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 329
Db         68 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 127
Qy        330 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 389
Db        128 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 187
Qy        390 GGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 449
Db        188 GGCAGGAAATCCACCCCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 247
Qy        450 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 509
Db        248 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 307
Qy        510 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 569
Db        308 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 367
Qy        570 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 629
```

```
Qy    361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db    420 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 479
Qy    421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db    480 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 539
Qy    481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db    540 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 599
Qy    541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db    600 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 659
Qy    601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db    660 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 719
Qy    661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db    720 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 779
Qy    721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db    780 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 839
Qy    781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db    840 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 899
Qy    841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db    900 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 959
Qy    901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db    960 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1019
Qy    961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db   1020 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1079
Qy   1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db   1080 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1139
Qy   1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db   1140 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1199
Qy   1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db   1200 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1259
Qy   1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db   1260 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1319
Qy   1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db   1320 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1379
Qy   1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db   1380 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1439
Qy   1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db   1440 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1499
Qy   1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db   1500 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1559
Qy   1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db   1560 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1619
Qy   1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db   1620 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1679
Qy   1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db   1680 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1739
Qy   1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db   1740 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1799
Qy   1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db   1800 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1859
Qy   1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db   1860 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1919
Qy   1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db   1920 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1979
Qy   1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db   1980 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2039
Qy   1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db   2040 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2099
Qy   2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db   2100 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2159
Qy   2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db   2160 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2219
Qy   2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db   2220 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2279
Qy   2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db   2280 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2339
Qy   2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db   2340 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2399
Qy   2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db   2400 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2459
Qy   2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db   2460 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2519
Qy   2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db   2520 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2579
Qy   2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
```

```
Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
Db       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAA 3816
```

RESULT 2
US-09-972-115A-5
; Sequence 5, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 4275
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (284)..(3781)
; OTHER INFORMATION:
US-09-972-115A-5

Query Match 99.4%; Score 3791.4; DB 4; Length 4275;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
Db         61 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAGGGGCGCGCG 120

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
Db        121 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 179

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
Db        180 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 239

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
Db        240 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 299

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db        300 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 359

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db        360 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 419
```

```
Db        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
```

```
Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Qy       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Qy       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Qy       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Qy       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Qy       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 19:31:32 ; Search time 249.12 Seconds
(without alignments)
8500.686 Million cell updates/sec

Title: US-09-843-159B-2
Perfect score: 3816
Sequence: 1 cgcgctgctccgcccgccgc.....cctttgctgaaaaaaaaaaa 3816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3816	100.0	3816	4	US-09-696-668-2	Sequence 2, Appli
	2	3791.4	99.4	4275	4	US-09-972-115A-5	Sequence 5, Appli
	3	3594.8	94.2	4512	4	US-09-350-982C-3	Sequence 3, Appli
	4	3488.8	91.4	3498	4	US-09-350-982C-4	Sequence 4, Appli
	5	3393.4	88.9	3797	4	US-09-696-668-1	Sequence 1, Appli
	6	3387.6	88.8	3394	4	US-09-427-154-1	Sequence 1, Appli
	7	3341	87.6	4297	4	US-09-972-115A-3	Sequence 3, Appli
	8	3039	79.6	4493	4	US-09-972-115A-1	Sequence 1, Appli
	9	1885	49.4	4134	3	US-09-196-387-1	Sequence 1, Appli
	10	1885	49.4	4134	4	US-09-841-835-1	Sequence 1, Appli
	11	1885	49.4	4134	4	US-09-972-115A-7	Sequence 7, Appli
	12	1521.2	39.9	4491	3	US-09-196-387-7	Sequence 7, Appli
	13	1521.2	39.9	4491	4	US-09-841-835-7	Sequence 7, Appli
	14	1369.4	35.9	4657	3	US-09-196-387-9	Sequence 9, Appli
	15	1369.4	35.9	4657	4	US-09-841-835-9	Sequence 9, Appli
c	16	518.4	13.6	520	4	US-09-972-115A-10	Sequence 10, Appl
	17	467.6	12.3	1069	4	US-09-833-381-841	Sequence 841, App
	18	461	12.1	465	4	US-09-833-381-1153	Sequence 1153, Ap
	19	380.4	10.0	384	4	US-09-972-115A-9	Sequence 9, Appli
	20	121.2	3.2	523	3	US-09-196-387-12	Sequence 12, Appl
	21	121.2	3.2	523	4	US-09-841-835-12	Sequence 12, Appl
	22	97.4	2.6	14770	4	US-09-220-132-30	Sequence 30, Appl
	23	93	2.4	5235	1	US-09-031-485-35	Sequence 35, Appl
c	24	93	2.4	5235	1	US-09-031-485-36	Sequence 36, Appl
	25	93	2.4	5235	1	US-08-847-429A-35	Sequence 35, Appl
c	26	93	2.4	5235	1	US-08-847-429A-36	Sequence 36, Appl
	27	93	2.4	5235	3	US-09-065-474-35	Sequence 35, Appl
c	28	93	2.4	5235	3	US-09-065-474-36	Sequence 36, Appl
	29	93	2.4	5235	4	US-09-557-034-35	Sequence 35, Appl
c	30	93	2.4	5235	4	US-09-557-034-36	Sequence 36, Appl
	31	93	2.4	5503	1	US-09-031-485-32	Sequence 32, Appl
c	32	93	2.4	5503	1	US-09-031-485-34	Sequence 34, Appl
	33	93	2.4	5503	1	US-08-847-429A-32	Sequence 32, Appl
c	34	93	2.4	5503	1	US-08-847-429A-34	Sequence 34, Appl
	35	93	2.4	5503	3	US-09-065-474-32	Sequence 32, Appl
c	36	93	2.4	5503	3	US-09-065-474-34	Sequence 34, Appl
	37	93	2.4	5503	4	US-09-557-034-32	Sequence 32, Appl
c	38	93	2.4	5503	4	US-09-557-034-34	Sequence 34, Appl
	39	83.4	2.2	1056	3	US-09-065-474-138	Sequence 138, App
c	40	83.4	2.2	1056	3	US-09-065-474-140	Sequence 140, App
	41	83.4	2.2	1056	4	US-09-557-034-138	Sequence 138, App
c	42	83.4	2.2	1056	4	US-09-557-034-140	Sequence 140, App
	43	82.6	2.2	909	1	US-09-031-485-25	Sequence 25, Appl
c	44	82.6	2.2	909	1	US-09-031-485-26	Sequence 26, Appl
	45	82.6	2.2	909	1	US-08-847-429A-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-696-668-2
; Sequence 2, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Chan, Eva
; APPLICANT: Xu, Xiang
; APPLICANT: Huang, Betty
; APPLICANT: Ossovskaya, Valeria
; TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS
; FILE REFERENCE: A-68292-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/696,668
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/427,154
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3816
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-696-668-2

Query Match 100.0%; Score 3816; DB 4; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
CC   given in the specification (including its fragment or variant), a vector
CC   comprising the SID (RTM) polynucleotide, a recombinant host cell
CC   comprising the vector, a protein chip comprising the polypeptides and a
CC   record comprising all or part of the data, listed in the specification.
CC   The complex, polypeptides, polynucleotides and compounds are useful for
CC   preventing or treating metabolic disorders such as obesity or diabetes.
CC   The polynucleotides are useful as probes or primers. The complex is
CC   particularly useful for identifying selected interacting domains (SID
CC   (RTM)) for screening drugs that modulate the protein interaction, thus
CC   exhibiting the therapeutic effect. The present sequence represents a SID
CC   (prey) protein of the invention
XX
SQ   Sequence 308 AA;

  Query Match             25.4%;  Score 1467;  DB 6;  Length 308;
  Best Local Similarity   88.3%;  Pred. No. 5e-126;
  Matches  272;  Conservative   23;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         18 VQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVL 77
              | |||||||||||||||||||||||:|||  |||||||||||||||||||||||||||||
Db          1 VHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYTPLHEAAIKGKIDVCIVL 60

Qy         78 LQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVN 137
              |||||:| ||||||::|||||||||||||||||||||||:||||||||:||||||||||
Db         61 LQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAARSGNEEKLMALLTPLNVN 120

Qy        138 CHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELL 197
              |||||||||||||||||||||:|||||||||||||||||| ||||||||||||||||||
Db        121 CHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLVPLHNACSYGHYEVTELL 180

Qy        198 VKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLK 257
              :|||||||||||||||||||||||||||||||:||||||:||| |||:|:||||:|:
Db        181 LKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLVNCHGKSAVDMAPTPELR 240

Qy        258 ERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQIC 317
              ||| |||||||||||||||: ::|| |:||::||| ||:||||||| || :|||||:
Db        241 ERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHETALHCAVASLHPKRKQVT 300

Qy        318 ELLLRKGA 325
              ||||||||
Db        301 ELLLRKGA 308


RESULT 45
AAB94719
ID   AAB94719 standard; protein; 261 AA.
XX
AC   AAB94719;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:15733.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 15733; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 261 AA;

  Query Match             24.2%;  Score 1395;  DB 4;  Length 261;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-119;
  Matches  260;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        825 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 60

Qy        885 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 944
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         61 SGTILIDLSPDDKEFPSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 120

Qy        945 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 180

Qy       1005 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 240

Qy       1065 VNGLALAEYVIYRGEQAYPEY 1085
              |||||||||||||||||||||
Db        241 VNGLALAEYVIYRGEQAYPEY 261


Search completed: March 25, 2004, 16:41:19
Job time : 62.1795 secs
```

```
Qy        452 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 300

Qy        512 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 360

Qy        572 DI 573
              ||
Db        361 DI 362
```

RESULT 43
AAB48572
ID AAB48572 standard; protein; 331 AA.
XX
AC AAB48572;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human breast cancer MO-BC-203 polypeptide.
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; MO-BC-203.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR N-PSDB; AAC89703, AAC89704.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 1; Page 78-79; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 331 AA;

Query Match 29.6%; Score 1706; DB 4; Length 331;
Best Local Similarity 97.9%; Pred. No. 5.1e-148;
Matches 324; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy        197 LVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQL 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQL 60

Qy        257 KERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQI 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQI 120

Qy        317 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 376
              |||||:||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CELLLKKGSNINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 180

Qy        377 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 240

Qy        437 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 496
              |||:||||||||||||||||||||||||||||||||||||||||||||||||:||||||
Db        241 GDVKTVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKEKGGLVP 300

Qy        497 LHNACSYGHYEVAELLVKHGAVVNVADLWKF 527
              |||||||||||||||||||| || :||||
Db        301 LHNACSYGHYEVAELLVKHGAVFNVTELWKF 331
```

RESULT 44
ABU70889
ID ABU70889 standard; protein; 308 AA.
XX
AC ABU70889;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human adipocyte Selected Interacting domain, SID, #520.
XX
KW Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KW antidiabetic; protein-protein interaction; diabetes;
KW yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS Homo sapiens.
XX
PN WO200286122-A2.
XX
PD 31-OCT-2002.
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Daviet L;
XX
DR WPI; 2003-103412/09.
DR N-PSDB; ACA57433.
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
PS Claim 6; Page 279; 382pp; English.
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM)) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences

PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR N-PSDB; ABK48634.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 138-140; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC protein sequence presented is the truncated human tankyrase #1. Note:
CC Residues 1-640 are the same as the wild-type human tankyrase sequence
CC (AAU79537)
XX
SQ Sequence 673 AA;

Query Match 33.3%; Score 1918.5; DB 5; Length 673;
Best Local Similarity 80.9%; Pred. No. 4.3e-167;
Matches 364; Conservative 42; Mismatches 43; Indels 1; Gaps 1;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:||| ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||:| |||||::|||||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||||:|||||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||:|||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:|||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:|
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYG 644
```

Qy 421 LGNSEADRQLLEAAKAGDVETVKKLCTVQS 450
|| | || |: |: | |::: : |
Db 645 SDPSIISLQGFTAAQMGN-EAVQQILSGHS 673

RESULT 42
AAB66300
ID AAB66300 standard; protein; 362 AA.
XX
AC AAB66300;
XX
DT 05-APR-2001 (first entry)
XX
DE Human TRF1 TANK2 binding domain protein sequence SEQ ID NO: 160.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63973.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 6; Page 228-229; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 362 AA;

Query Match 32.8%; Score 1895; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 2.2e-165;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        212 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 FTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQ 60

Qy        272 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKT 120

Qy        332 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGC 180

Qy        392 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSV 240
```

```
  Query Match             33.7%;  Score 1942;  DB 4;  Length 385;
  Best Local Similarity   99.5%;  Pred. No. 1.1e-169;
  Matches  369;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         37 HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALD 96
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALD 60

Qy         97 LADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYN 156
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 LADPSAKAVLTGEYKKDELLESARXGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYN 120

Qy        157 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH 216
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLH 180

Qy        217 EAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREA 276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 EAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREA 240

Qy        277 DVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 300

Qy        337 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNII 396
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        301 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLHTCRLLLSYGCDPNII 360

Qy        397 SLQGFTALQMG 407
              ||||||||||| 
Db        361 SLQGFTALQMG 371
```

RESULT 40
AAY44403
ID AAY44403 standard; protein; 673 AA.
XX
AC AAY44403;
XX
DT 22-MAR-2000 (first entry)
XX
DE Human truncated tankyrase-1.
XX
KW Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW telomere length regulation; recombinant tankyrase; aging; skin atrophy;
KW macular degeneration; atherosclerosis; ataxia telangiectasia;
KW tankyrase modulator; rational drug design.
XX
OS Homo sapiens.
XX
PN WO9964606-A1.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-US012968.
XX
PR 10-JUN-1998; 98US-00095225.
PR 17-AUG-1998; 98US-00135233.
PR 19-NOV-1998; 98US-00196387.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2000-116549/10.
DR N-PSDB; AAZ29632.
XX
PT New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT length, used to identify modulators, e.g. for inhibiting growth of
PT cancer.
XX
PS Disclosure; Page 121-123; 133pp; English.
XX
CC The present sequence is truncated human tankyrase, a TRF1 (telomeric
CC repeat binding factor-1) binding protein. This comprises 1-640 residues
CC of human tankyrase. Tankyrase is involved in regulation of telomere
CC length through TRF1 and may directly modulate the effect of TRF1.
CC Recombinant tankyrase, or its fragments are used to identify specific
CC modulators which are potential drugs for countering telomere shortening
CC associated with aging (e.g. atrophy of skin, macular degeneration or
CC atherosclerosis) or diseases like ataxia telangiectasia. They are also
CC used to raise specific antibodies (used as immunoassay reagents and as
CC modulators of tankyrase activity) and in rational drug design
XX
SQ Sequence 673 AA;

```
  Query Match             33.3%;  Score 1918.5;  DB 3;  Length 673;
  Best Local Similarity   80.9%;  Pred. No. 4.3e-167;
  Matches  364;  Conservative   42;  Mismatches   43;  Indels    1;  Gaps    1;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:|||  ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||:| |||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||||:|||||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::||  |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:|
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYG 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQS 450
                  |    |   ||: |: | |::: :  |
Db        645 SDPSIISLQGFTAAQMGN-EAVQQILSGHS 673
```

RESULT 41
AAU79538
ID AAU79538 standard; protein; 673 AA.
XX
AC AAU79538;
XX
DT 12-AUG-2002 (first entry)
XX
DE Truncated human tankyrase #1.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; cytostatic; ADP ribosylation;
KW telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia; enzyme.
XX
OS Homo sapiens.
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX

```
XX
DT   10-JUN-2003  (first entry)
XX
DE   Human adipocyte Selected Interacting domain, SID, #509.
XX
KW   Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KW   antidiabetic; protein-protein interaction; diabetes;
KW   yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS   Homo sapiens.
XX
PN   WO200286122-A2.
XX
PD   31-OCT-2002.
XX
PF   14-MAR-2002; 2002WO-EP003768.
XX
PR   14-MAR-2001; 2001US-0275734P.
XX
PA   (HYBR-) HYBRIGENICS.
XX
PI   Legrain P,  Daviet L;
XX
DR   WPI; 2003-103412/09.
DR   N-PSDB; ACA57422.
XX
PT   New complex between two interacting proteins in adipocyte cells, useful
PT   for identifying selected interacting domains that modulate protein
PT   interactions, or for preventing or treating metabolic disorders such as
PT   obesity or diabetes.
XX
PS   Claim 6; Page 276; 382pp; English.
XX
CC   The invention relates to a complex between two interacting proteins in
CC   adipocyte cells, given in the specification. The proteins are identified
CC   by selecting a bait protein from a known adipocyte marker and then
CC   performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC   members of an adipocyte cDNA library. The proteins are designated SID
CC   (RTM) (selected interacting domains) proteins. Also included are a
CC   polynucleotide encoding a polypeptide in the adipocyte cells, a
CC   recombinant host cell expressing at least one of the interacting
CC   polypeptides of the complex, selecting a modulating compound in adipocyte
CC   cells, a SID (RTM)) polypeptide comprising any of the 738 amino acid
CC   sequences given in the specification (including its fragment or variant),
CC   a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC   given in the specification (including its fragment or variant), a vector
CC   comprising the SID (RTM) polynucleotide, a recombinant host cell
CC   comprising the vector, a protein chip comprising the polypeptides and a
CC   record comprising all or part of the data, listed in the specification.
CC   The complex, polypeptides, polynucleotides and compounds are useful for
CC   preventing or treating metabolic disorders such as obesity or diabetes.
CC   The polynucleotides are useful as probes or primers. The complex is
CC   particularly useful for identifying selected interacting domains (SID
CC   (RTM)) for screening drugs that modulate the protein interaction, thus
CC   exhibiting the therapeutic effect. The present sequence represents a SID
CC   (prey) protein of the invention
XX
SQ   Sequence 373 AA;

  Query Match             34.0%;  Score 1962;  DB 6;  Length 373;
  Best Local Similarity   99.7%;  Pred. No. 1.5e-171;
  Matches  372;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 QNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKID 72
              ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 QNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKID 60

Qy         73 VCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLT 120

Qy        133 PLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYE 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 PLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYE 180

Qy        193 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 240

Qy        253 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 300

Qy        313 RKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 RKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 360

Qy        373 RAAYCGHLQTCRL 385
              |||||||||||||
Db        361 RAAYCGHLQTCRL 373


RESULT 39
AAB48571
ID   AAB48571 standard; protein; 385 AA.
XX
AC   AAB48571;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Human T cell leukaemia HOM-TALL1-5 polypeptide.
XX
KW   Human; cytostatic; gene therapy; vaccine; breast cancer;
KW   T cell leukaemia; cancer associated antigen; HOM-TALL1-5.
XX
OS   Homo sapiens.
XX
PN   WO200072021-A2.
XX
PD   30-NOV-2000.
XX
PF   24-MAY-2000; 2000WO-US014391.
XX
PR   26-MAY-1999;   99US-00320092.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Kuimov A,  Lagarkova M,  Koroleva E,  Turetskaya R,  Vdovichenko K;
PI   Mescheryakov A,  Litchinitser M,  Kuprash D,  Nedospasov S,  Tureci O;
PI   Sahin U,  Pfruendschuh M,  Old LJ,  Knuth A,  Jager E;
XX
DR   WPI; 2001-032067/04.
DR   N-PSDB; AAC89701, AAC89702.
XX
PT   Diagnosing disorder characterized by expression of human cancer
PT   associated antigen precursor, involves detecting interaction of an agent
PT   with novel NA Group 1 nucleic acid molecule encoding the antigen
PT   precursor.
XX
PS   Example 1; Page 75-76; 87pp; English.
XX
CC   The present sequence is given in a specification relating to nucleic
CC   acids and encoded polypeptides which are cancer associated antigens
CC   expressed in patients afflicted with breast cancer and/or T cell
CC   leukaemia. A disorder characterised by expression of human cancer
CC   associated antigen may be diagnosed by contacting a biological sample
CC   with an agent that specifically binds to the nucleic acid that encodes
CC   the antigen, complexed with human leukocyte antigen molecule, and
CC   determining the interaction between the agent and the nucleic acid.
CC   Cancer associated antigens, the nucleotides encoding them, antibodies
CC   against them and the pharmaceutical compositions comprising them are
CC   useful for diagnosing, monitoring and treating the diseases characterised
CC   by the expression of one or more cancer associated antigens
XX
SQ   Sequence 385 AA;
```

```
CC   polypeptide of the invention
XX
SQ   Sequence 415 AA;

  Query Match             37.5%;  Score 2165;  DB 4;  Length 415;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-190;
  Matches  413;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        208 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGH 60

Qy        268 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANI 120

Qy        328 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLL 180

Qy        388 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCT 240

Qy        448 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYE 300

Qy        508 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVK 360

Qy        568 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 DGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA 413


RESULT 37
AAB66301
ID   AAB66301 standard; protein; 1099 AA.
XX
AC   AAB66301;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 expression plasmid protein sequence SEQ ID NO: 178.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63990.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 7; Page 236-239; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1099 AA;

  Query Match             35.6%;  Score 2053.5;  DB 4;  Length 1099;
  Best Local Similarity   62.1%;  Pred. No. 3.6e-179;
  Matches  459;  Conservative   38;  Mismatches  115;  Indels  127;  Gaps   17;

Qy        404 LQMGNENVQQLLQEGISLG---NSEADRQLLEAAKAGDVETV-KKLCTVQSVNCRDIEGR 459
              |    : |: ::::   ||       |::  |  :   :|| : ||:  |:  | |
Db        437 LSRNKDEVKAMIEK---LGGKLTGTANKASLCISTKKEVEKMNKKMEEVKEANIR----- 488

Qy        460 QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
                       | |  ||   || |  |             |     :| :|   || |
Db        489 --------------VVSEDFLQ---DVSASTK----------SLQELFLAHILSPWGAEV 521

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 579
                    |  |:  | :||          ||  :||:: |    : :   :  ::  |:|
Db        522 ------KAEPVEVVAPRGK---------SGAALSKKSK-GQVKEEGINKSEKRMKLTLKG 565

Qy        580 DAAL-----------------LDAAKKGCLARVK--------KLSSPDNVN---CRDTQ 610
              ||:                 : :|  | :  ||        :|   |  |     :
Db        566 GAAVDPDSGLEHSAHVLEKGGKVFSATLGLVDIVKGTNSYYKLQLLEDDKENRYWIFRSW 625

Qy        611 GRHSTPLHLAAGYNNLE---VAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIK 667
              ||  | :    | | ||     |  ::|   :  :  |      |   |
Db        626 GRVGTVI----GSNKLEQMPSKEDAIEHFMKLYEEKTGNAWHSKNFTKYPK--------- 672

Qy        668 YNACVNATDKWAFTPLH------EAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 721
                        | ||       | | |  |              :|  |  ||||||||||
Db        673 -----------KFYPLEIDYGQDEEAVKKLT----------VNPGTKRVEGQTPLDLVSA 711

Qy        722 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 781
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 771

Qy        782 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 841
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 831

Qy        842 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 901
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 891

Qy        902 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 961
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER 951

Qy        962 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1021
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDR 1011

Qy       1022 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1081
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQA 1071

Qy       1082 YPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||
Db       1072 YPEYLITYQIMRPEGMVDG 1090


RESULT 38
ABU70878
ID   ABU70878 standard; protein; 373 AA.
XX
AC   ABU70878;
```

XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63925.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 156-157; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 522 AA;

Query Match 41.0%; Score 2367; DB 4; Length 522;
Best Local Similarity 99.8%; Pred. No. 1.2e-208;
Matches 452; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db         70 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 129

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 189

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 249

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 309

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 369

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 429

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 489

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 453
              |||||||||||||||||||||||||||||||||
Db        490 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNC 522
```

RESULT 36
ABB11320
ID ABB11320 standard; peptide; 415 AA.
XX
AC ABB11320;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human Grb7 effector protein homologue, SEQ ID NO:1690.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA08564.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 20; Page 165-166; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human

```
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .22
FT                   /label= Ankyrin_repeat_1
FT   Region          24. .55
FT                   /label= Ankyrin_repeat_2
FT   Region          57. .87
FT                   /label= Ankyrin_repeat_3
FT   Region          89. .113
FT                   /label= Ankyrin_repeat_4
FT   Region          115. .139
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 127
FT                   /note= "TH1 start"
FT   Region          147. .176
FT                   /label= Ankyrin_repeat_6
FT   Region          178. .208
FT                   /label= Ankyrin_repeat_7
FT   Region          210. .241
FT                   /label= Ankyrin_repeat_8
FT   Region          243. .281
FT                   /label= Ankyrin_repeat_9
FT   Region          283. .332
FT                   /label= Ankyrin_repeat_10
FT   Region          334. .365
FT                   /label= Ankyrin_repeat_11
FT   Region          367. .398
FT                   /label= Ankyrin_repeat_12
FT   Region          400. .424
FT                   /label= Ankyrin_repeat_13
FT   Region          426. .458
FT                   /label= Ankyrin_repeat_14
FT   Region          460. .491
FT                   /label= Ankyrin_repeat_15
FT   Region          493. .524
FT                   /label= Ankyrin_repeat_16
FT   Region          526. .557
FT                   /label= Ankyrin_repeat_17
FT   Region          559. .584
FT                   /label= Ankyrin_repeat_18
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-C; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform TH-1 variant TaHo C which
CC   contains a C-terminus deletion and a 28 amino acid insertion due to a
CC   frame shift described in the disclosure of the invention
XX
SQ   Sequence 584 AA;

  Query Match             50.5%;  Score 2916;  DB 6;  Length 584;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-259;
  Matches  555;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKK 555
              |||||||||||||||
Db        541 ICKLLLQHGADPTKK 555


RESULT 35
AAB66287
ID   AAB66287 standard; protein; 522 AA.
XX
AC   AAB66287;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone protein sequence SEQ ID NO: 99.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
```

```
Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:||||||||||||||||||||:||||||||:||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH-- 658
              |:|:||||||||:|||||||||||||||||||:||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCL 884

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 716
                |  |    |     |    |||| ||     ::   || ||||   ::: |  ||
Db        885 ARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942
```

RESULT 33
AAU79539
ID AAU79539 standard; protein; 949 AA.
XX
AC AAU79539;
XX
DT 12-AUG-2002 (first entry)
XX
DE Truncated human tankyrase #2.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; cytostatic; ADP ribosylation;
KW telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia; enzyme.
XX
OS Homo sapiens.
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR N-PSDB; ABK48635.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 146-148; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC protein sequence presented is the truncated human tankyrase #2. Note:
CC Residues 1-881 are the same as the wild-type human tankyrase sequence
CC (AAU79537)
XX
SQ Sequence 949 AA;

Query Match 53.6%; Score 3093; DB 5; Length 949;
Best Local Similarity 81.5%; Pred. No. 3.4e-275;
Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:|||  ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||:| ||||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:||||||||||||||||||||||||||:||||||||||||||||||  ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEPLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: ||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:||||||:||| |||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:||||||||||||||||||||:||||||||:||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH-- 658
              |:|:||||||||:|||||||||||||||||||:||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGGCL 884

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPL 716
                |  |    |     |    |||| ||     ::   || ||||   ::: |  ||
Db        885 ARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPL 942
```

RESULT 34
ABG73734
ID ABG73734 standard; protein; 584 AA.
XX
AC ABG73734;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 1 TH-1 variant TaHo C.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo C;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP; variant;
KW mutant.

```
Qy        355 KHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQL 414

Db        321 ------------------------------------------------------------ 320

Qy        415 LQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 474
                                          ||||||||||||||||||||||||||||||
Db        321 ----------------------------KLCTVQSVNCRDIEGRQSTPLHFAAGYNRVS 351

Qy        475 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 VVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA 411

Qy        535 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLAR 471

Qy        595 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 VKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAA 531

Qy        655 SYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQT 714
              |||
Db        532 SYG---------------------------------------------------------- 534

Qy        715 PLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASS 774

Db        535 ------------------------------------------------------------ 534

Qy        775 LDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQIT 834
                                                                        ||
Db        535 ------------------------------------------------------------IT 536

Qy        835 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 LDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSP 596

Qy        895 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        597 DDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEEN 656

Qy        955 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        657 HNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTG 716

Qy       1015 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        717 CPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYV 776

Qy       1075 IYRGEQAYPEYLITYQIMRPEGMVDG 1100
              |||||||||||||||||||||||||
Db        777 IYRGEQAYPEYLITYQIMRPEGMVDG 802
```

```
RESULT 32
AAY44404
ID   AAY44404 standard; protein; 949 AA.
XX
AC   AAY44404;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human truncated tankyrase-2.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   telomere length regulation; recombinant tankyrase; aging; skin atrophy;
KW   macular degeneration; atherosclerosis; ataxia telangiectasia;
KW   tankyrase modulator; rational drug design.
XX
OS   Homo sapiens.
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   N-PSDB; AAZ29633.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Disclosure; Page 125-129; 133pp; English.
XX
CC   The present sequence is truncated human tankyrase, a TRF1 (telomeric
CC   repeat binding factor-1) binding protein. This comprises 1-881 residues
CC   of human tankyrase. Tankyrase is involved in regulation of telomere
CC   length through TRF1 and may directly modulate the effect of TRF1.
CC   Recombinant tankyrase, or its fragments are used to identify specific
CC   modulators which are potential drugs for countering telomere shortening
CC   associated with aging (e.g. atrophy of skin, macular degeneration or
CC   atherosclerosis) or diseases like ataxia telangiectasia. They are also
CC   used to raise specific antibodies (used as immunoassay reagents and as
CC   modulators of tankyrase activity) and in rational drug design
XX
SQ   Sequence 949 AA;
```

Query Match 53.6%; Score 3093; DB 3; Length 949;
Best Local Similarity 81.5%; Pred. No. 3.4e-275;
Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||:||| ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||:| ||||||:||||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||||||||:|||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||:|||||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  ||||||||||||| ||:|||||||| |||||  |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||::||:  :|||||:|||  ||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764
```

```
Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              |||:|||||  |||  :|:||:|:|:|||  ||  |  |:  ||:|  ||  :   |:|||  ||  |  :|
Db        306 NCHSKSAIDAAPTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDT 365

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||  |   ||   ||||:  |||  |||:  :|||  |  ||||||:|:|   |  |  :||::|   |||
Db        366 PLHLAVVSPDGKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKV 425

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||:||||  |||  |      |  ||||||   |  ||:||:|  ||  |:  :::|  :||:
Db        426 NALDSLGQTPLHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK---- 479

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                  :::  |||||||||::||:::       ||||||::||  |||||||||:|||  ||::
Db        480 -NPPDSETHLLEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQF 538

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
              ||:|||:|:|  ||||||||||||||||||  ||||||||  |||:|||||||||||||||
Db        539 LLEHGAEVYAADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGK 598

Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
              |:|||||:|||||  ||||||  ||  ||||:  |  |:  :|||  :|||||||||  ||||::|
Db        599 YDICKLLLKHGADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRL 658

Qy        599 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658
              :|:::||||  |||:||||||||||||  |  |||||::|||||||||||||||||:||||
Db        659 VTPESINCRDAQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGH 718

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
              :|:||||||:   |||||||  ||||||||||||||||:||||||||  :||||||||::|
Db        719 LDIAALLIKHKTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIEL 778

Qy        719 VSADDVSALLTAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAA 772
              :||||  ||  ||  |    ||  |      : :   :| |||  |      ||  ||  ||
Db        779 ATADDVKCLLLQDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPT 837

Qy        773 SSLDNLSGSFSELSSV-----VSSS-----------GTEGASSLEKKEVPGVD--FSIT 813
              :    |    | : ||     :|||           | ||:|| :   :|  |   :::
Db        838 TETVLLPTGASMILSVPVPLPLSSSTRISPAQGAEANGAEGSSSDDL--LPDADTITNVS 895

Qy        814 QFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 873
              |: :  | ||:::|||||||||:| |||| :||::|::|||  |||::||: :| |
Db        896 GFLSSQQLHHLIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS---- 951

Qy        874 LNPYLTLNTSGSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYN 927
                    |:|  |       |:|:|| ||||||  :||||||:|:|||||  | |||  | |||
Db        952 --------TTGIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYN 1003

Qy        928 ILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGG 987
              |:::|||  |:||||||  |||:|::|||   :|||||||||||:|||: :|||||||||||
Db       1004 IIRVQKVQNRKLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGG 1063

Qy        988 MFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAM 1047
              ||||||||||:|||||||||||||| |||  |||:|||:|  ||||  |||  |||||||:|||
Db       1064 MFGAGIYFAEHSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAM 1123

Qy       1048 KMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||:|||||| ||||  ||  ||||:|||||:||||||||||::|:    |
Db       1124 KMAHAPPGHHSVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176


RESULT 31
AAO20512
ID   AAO20512 standard; protein; 802 AA.
XX
AC   AAO20512;
XX
DT   27-JUN-2002  (first entry)
XX
DE   Protein of APP related human homologue hCP33787.
XX
KW   Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW   amyloid precursor protein; tissue-specific expression control; human APP;
KW   APP pathway modulator; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200226820-A2.
XX
PD   04-APR-2002.
XX
PF   01-OCT-2001; 2001WO-EP011345.
XX
PR   29-SEP-2000; 2000US-0236893P.
PR   14-JUN-2001; 2001US-0298309P.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI   Cohen D,  Dengler UJ,  Finelli AL,  Freuler F,  Konsolaki M;
PI   Reinhardt MWHM,  Zusman S;
XX
DR   WPI; 2002-315796/35.
DR   N-PSDB; AAK99406.
XX
PT   New transgenic fly, containing DNA encoding an Abeta portion of human
PT   APP, useful for identifying agents which modulate the APP pathway and
PT   which can be used to treat Alzheimer's disease.
XX
PS   Example 4; Page 113-114; 129pp; English.
XX
CC   The invention relates to a transgenic fly whose genome comprises DNA
CC   encoding a polypeptide having the Abeta portion of human amyloid
CC   precursor protein (APP), fused to a signal sequence. The DNA sequence
CC   encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC   the specification. The DNA sequence is operably linked to a tissue-
CC   specific expression control sequence. Expression of the sequence gives
CC   the fly an altered phenotype. The purpose of the invention is for
CC   identifying agents that inhibit or promote the expression and/or function
CC   of genes or encoded polypeptides which modify the APP pathway. The agent
CC   is a compound, triple helix DNA, antisense oligonucleotide, double
CC   stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC   to treat conditions such as Alzheimer's disease. The agent can be used as
CC   an APP pathway modulator or in gene therapy. This sequence represents the
CC   protein of the APP related human homologue hCP33787
XX
SQ   Sequence 802 AA;

  Query Match           62.3%;  Score 3596;  DB 5;  Length 802;
  Best Local Similarity  65.1%;  Pred. No. 0;
  Matches  720;  Conservative    5;  Mismatches    5;  Indels  376;  Gaps    5;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGE-----YKKDEL 115
              |||||||||||||||||||||||||||||||||||||||||||||||||:     | :  :
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGKSVYSGYSRKPV 186

Qy        116 LES-ARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAK 174
                  : |||||||||||||||||||||||
Db        187 KNNLARSGNEEKMMALLTPLNVNCHASDGR---------------------------- 216

Qy        175 DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 234
                                          |||||||||||||||||||||||||||||||
Db        217 ------------------------KHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYG 252

Qy        235 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 ADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKH 312

Qy        295 PQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVV 354
              ||||||||
Db        313 PQTHETAL---------------------------------------------------- 320
```

```
  Best Local Similarity   65.3%;  Pred. No. 0;
  Matches  740;  Conservative   156;  Mismatches   182;  Indels    55;  Gaps    12;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |:||::|||:|| :|||:|| |:||| |||| |||||||| |||: || ||  ||||||
Db         66 GYGRREVVEFLLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYT 125

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||| |||:|||: |||||  ||||:: :| |:||| : : |||||:|||||:||
Db        126 PLHEAASKGKVDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAAR 185

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              || |::::|||||||||||||||:||||||||||||: ||::|| :|||||||||| ||
Db        186 SGAEDRLLALLTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLV 245

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||::||:||:: || ||| ||| |||||||||:|||||||||| |||||||
Db        246 PLHNACSYGHFDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLL 305

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              |||:|||| ||| :|:||:|:|:||| || | |: ||:| || :  |:||| || | :|
Db        306 NCHSKSAIDAAPTRELRERIAFEYKGHCLLDACRKCDVSRAKKLVCAEIVNFVHPYTGDT 365

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              || |  ||  |||: ||| |||: :||| | |||||:|:|  | | :||::|  |||
Db        366 PLHLAVVSPDGKRKQLMELLTRKGSLLNEKNKAFLTPLHLAAELLHYDAMEVLLKQGAKV 425

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||:|||| ||| |   |  ||||||  | ||:||:| || |: :::| :|::
Db        426 NALDSLGQTPLHRCA--RDEQAVRLLLSYAADTNIVSLEGLTAAQLASDSVLKLLK---- 479

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQ--SVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                  :::  |||||||||::||:::      ||||||::|| |||||||||:||| ||::
Db        480 -NPPDSETHLLEAAKAGDLDTVRRIVLNNPISVNCRDLDGRHSTPLHFAAGFNRVPVVQF 538

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
              ||:|||:|:| |||||||||||||||||| |||||||| |||:|||||||||||||||||
Db        539 LLEHGAEVYAADKGGLVPLHNACSYGHYEVTELLVKHGANVNVSDLWKFTPLHEAAAKGK 598

Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
              |:|||||:|||| |||||| || |||: | |: :|||| :|||||||| ||||::|
Db        599 YDICKLLLKHGADPMKKNRDGATPADLVKESDHDVAELLRGPSALLDAAKKGNLARVQRL 658

Qy        599 SSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658
              :|:::|||| |||:||||||||||| | |||||::|||||||||||||||||:|||
Db        659 VTPESINCRDAQGRNSTPLHLAAGYNNFECAEYLLENGADVNAQDKGGLIPLHNASSYGH 718

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
              :|:|||||:   ||||||| |||||||||||||||||:|||||||  :||||||||::|
Db        719 LDIAALLIKHKTVVNATDKWGFTPLHEAAQKGRTQLCSLLLAHGADAYMKNQEGQTPIEL 778

Qy        719 VSADDVSALLTAAMPPS----ALPSCYKPQVLNGVRSPGATADALSSGPSSPSS--LSAA 772
              :||||  ||  ||  |   || |     : :   :| ||| |     || ||  ||
Db        779 ATADDVKCLLQDAMATSLSQQAL-SASTQSLTSSSPAPDATAAAAPGTSSSSSSAILSPT 837

Qy        773 SSLDNLSGSFSELSSV-----VSSS-----------GTEGASSLEKKEVPGVD--FSIT 813
              :    |    | : ||     :|||           | ||:|| :   :|  |   :::
Db        838 TETVLLPTGASMILSVPVPLPLSSSTRISPAQGAEANGAEGSSSDDL--LPDADTITNVS 895

Qy        814 QFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQG 873
              |: :  | ||:::|||||||||:| |||| :||::|::||| |||::||: : | |
Db        896 GFLSSQQLHHLIELFEREQITLDILAEMGHDDLKQVGVSAYGFRHKILKGIAQLRS---- 951

Qy        874 LNPYLTLNTSGSG------TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYN 927
                    |:| |      |:|:|| |||||| :||||||:|:|||| | ||| | |||
Db        952 --------TTGIGNNVNLCTLLVDLLPDDKEFVAVEEEMQATIREHRDNGQAGGYFTRYN 1003

Qy        928 ILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGG 987
              |:::||| |:|||||| |||:|::||   :|||||||||||:||: :||||||||||
Db       1004 IIRVQKVQNRKLWERYAHRRQEIAEENFLQSNERMLFHGSPFINAIVQRGFDERHAYIGG 1063

Qy        988 MFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAM 1047
              ||||||||||:|||||||||||||| ||| |||:|||:| |||| ||| |||||||:|||

Db       1064 MFGAGIYFAEHSSKSNQYVYGIGGGIGCPSHKDKSCYVCPRQLLLCRVALGKSFLQYSAM 1123

Qy       1048 KMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||:|||||| |||| || ||||:|||||:||||||||::|:    |
Db       1124 KMAHAPPGHHSVVGRPSAGGLHFAEYVVYRGEQSYPEYLITYQIVKPDDSSSG 1176


RESULT 30
AAB66297
ID   AAB66297 standard; protein; 1181 AA.
XX
AC   AAB66297;
XX
DT   05-APR-2001  (first entry)
XX
DE   Drosophila tankyrase SEQ ID NO: 139.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Drosophila melanogaster.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63955.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 2; Page 215-218; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1181 AA;

  Query Match             66.5%;  Score 3835.5;  DB 4;  Length 1181;
  Best Local Similarity   65.3%;  Pred. No. 0;
  Matches  740;  Conservative   156;  Mismatches   182;  Indels    55;  Gaps    12;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |:||::|||:|| :|||:|| |:||| |||| |||||||| |||: || ||  ||||||
Db         66 GYGRREVVEFLLNSGASIQACDEGGLHPLHNCCSFGHAEVVRLLLKAGASPNTTDNWNYT 125

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||| |||:|||: |||||  ||||:: :| |:||| : : |||||:|||||:||
Db        126 PLHEAASKGKVDVCLALLQHGANHTIRNSEQKTPLELADEATRPVLTGEYRKDELLEAAR 185

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              || |::::|||||||||||||||:||||||||||||: ||::|| :|||||||||| ||
Db        186 SGAEDRLLALLTPLNVNCHASDGRRSTPLHLAAGYNRIGIVEILLANGADVHAKDKGGLV 245

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||::||:||:: || ||| ||| |||||||||:|||||||||| |||||||
Db        246 PLHNACSYGHFDVTKLLIQAGANVNANDLWAFTPLHEAASKSRVEVCSLLLSRGADPTLL 305
```

XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63918.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 150-152; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 756 AA;

Query Match 68.6%; Score 3959; DB 4; Length 756;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 756; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        345 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL 60

Qy        405 QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPL 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPL 120

Qy        465 HFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 HFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADL 180

Qy        525 WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALL 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALL 240

Qy        585 DAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDK 300

Qy        645 GGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGAD 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGAD 360

Qy        705 PTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPS 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPS 420

Qy        765 SPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL 480

Qy        825 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSG 540

Qy        885 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYT 600

Qy        945 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 HRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQ 660

Qy       1005 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 YVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPS 720

Qy       1065 VNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||||||||||||||||
Db        721 VNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 756
```

RESULT 29
ABB60894
ID ABB60894 standard; protein; 1181 AA.
XX
AC ABB60894;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9474.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL04997.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 9474; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1181 AA;

Query Match 66.5%; Score 3835.5; DB 4; Length 1181;

```
Qy        733 PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSS 792
              ||||||||||||||||||||||||||     : :    : |  :   ||||||||||
Db        541 PPSALPSCYKPQVLNGVRSPGATADALFQVHLAHQAFLQPAVLTTI-WEFSELSSVVSSS 599

Qy        793 GTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 GTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN 659

Qy        853 AYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVRE 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 AYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVRE 719

Qy        913 HRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 HRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNA 779

Qy        973 IIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLF 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 IIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLF 839

Qy       1033 CRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIM 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 CRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIM 899

Qy       1093 RPEGMVDG 1100
              ||||||||
Db        900 RPEGMVDG 907
```

RESULT 27
AAB66285
ID AAB66285 standard; protein; 784 AA.
XX
AC AAB66285;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone protein sequence SEQ ID NO: 89.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63917.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 144-146; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 784 AA;

Query Match 71.1%; Score 4102; DB 4; Length 784;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 784; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        317 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAY 60

Qy        377 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKA 120

Qy        437 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVP 180

Qy        497 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 240

Qy        557 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP 300

Qy        617 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATD 360

Qy        677 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA 420

Qy        737 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEG 480

Qy        797 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGH 540

Qy        857 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 RHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDG 600

Qy        917 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHK 660

Qy        977 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVT 720

Qy       1037 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 LGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEG 780

Qy       1097 MVDG 1100
              ||||
Db        781 MVDG 784
```

RESULT 28
AAB66286
ID AAB66286 standard; protein; 756 AA.
XX
AC AAB66286;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone protein sequence SEQ ID NO: 91.

```
Db        427 NALDSLGQTSLHRAAHCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQG--- 483

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
                                 ||||||||||||||||||||||||||||||||||||||||
Db        484 ------------------ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 524

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 584

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||||||||||||||||||||||||||| | || ||||||||||||||||||||||
Db        585 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTQIGKLLIGDAALLDAAKKGCLARVKKLSS 644

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 704

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        705 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 764

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||
Db        765 ADDVSALLTAAM---------------------------------------------------- 776

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Db        777 ---------------------------------------------------------- 776

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
                                                  ||||||||||||||
Db        777 ----------------------------------------------GTILIDLSPDDKEFQ 791

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db        792 SVEEEMQSTVREHRDGGHAGGVFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 851

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        852 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPIHKD 911

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 971

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||:||||||||
Db        972 AYPEYLITYQIVRPEGMVDG 991
```

RESULT 26
AAB48574
ID AAB48574 standard; protein; 907 AA.
XX
AC AAB48574;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human breast cancer MO-BC-203 polypeptide.
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; MO-BC-203.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR N-PSDB; AAC89706.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 8; Page 84-86; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 907 AA;

Query Match 80.2%; Score 4624.5; DB 4; Length 907;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 885; Conservative 6; Mismatches 16; Indels 1; Gaps 1;

```
Qy        193 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAP 60

Qy        253 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPK 120

Qy        313 RKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 372
              |||||||||:||:||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RKQICELLLKKGSNINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLH 180

Qy        373 RAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLE 240

Qy        433 AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKG 300

Qy        493 GLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADP 552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADP 360

Qy        553 TKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGR 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGR 420

Qy        613 HSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 HSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACV 480

Qy        673 NATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM 540
```

```
Qy        482 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 541
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        646 HGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 705

Qy        542 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 601
              |||||||||||||| : |   ||||||| ||||| ||||| ||||||||||||||    |
Db        706 CKLLLQHGADPTKKKKKGXIXLDLVKDGXTDIQDXLRGDAVLLDAAKKGCLARVKKXXFP 765

Qy        602 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 661
              |||||||||||||||| ||
Db        766 DNVNCRDTQGRHSTPLHL-AGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 824

Qy        662 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 721

Db        825 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 884

Qy        722 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 781
              |||||||  |||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        885 XXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 944

Qy        782 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 841
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        945 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVLVEM 1004

Qy        842 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 901
              |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1005 GHKELKEIGINAYGHRHKLIKSFERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1064

Qy        902 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK---KLWERYTHRRKEVSEENHNHA 958
              |||||||||||||||||||||||||||||||||:   :  ||||||||||||||||||||
Db       1065 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNRAKIRHEERYTHRRKEVSEENHNHA 1124

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1125 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1184

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1185 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1244

Qy       1079 EQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||
Db       1245 EQAYPEYLITYQIMRPEGMVDG 1266
```

```
RESULT 25
AAB47023
ID   AAB47023 standard; protein; 991 AA.
XX
AC   AAB47023;
XX
DT   29-MAR-2001  (first entry)
XX
DE   Mouse SPANK.
XX
KW   SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW   poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW   ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW   adipocyte; insulin signalling pathway; hyperlipidaemia;
KW   glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW   cardiac insufficiency; coronary insufficiency; stroke;
KW   high blood pressure; non-insulin dependent diabetes; hypertension;
KW   hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 483
FT                   /note= "Encoded by G"
FT   Misc-difference 614. .623
FT                   /note= "Encoded by G"
FT   Misc-difference 776
FT                   /note= "Encoded by ATGCC"
XX
PN   WO200077225-A1.
XX
PD   21-DEC-2000.
XX
PF   09-JUN-2000; 2000WO-US015926.
XX
PR   11-JUN-1999;   99US-0138957P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Chi N,  Lodish HF;
XX
DR   WPI; 2001-091404/10.
DR   N-PSDB; AAC85295.
XX
PT   New insulin signaling protein SPANK, useful for reducing body mass,
PT   glucose intolerance or insulin resistance and for preventing or treating
PT   obesity-related and muscle-related diseases.
XX
PS   Claim 3; Fig 5B; 65pp; English.
XX
CC   This sequence represents the mouse SPANK protein. The SPANK protein
CC   comprises 3 domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP
CC   (poly adenosine diphosphate-ribose polymerase) catalytic domain; and (c)
CC   an ANK domain composed of ankyrin repeats. SPANK is a cytosolic protein
CC   which can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC   aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC   perinuclear region of adipocytes. It is an effector in the insulin
CC   signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC   mass, reducing glucose intolerance or insulin resistance, for preventing
CC   or treating obesity-related diseases or disorders, such as obesity,
CC   cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC   atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC   dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC   also useful for preventing or treating muscle-related diseases or
CC   disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC   Antibodies immunospecific for SPANK are useful for detecting the presence
CC   of SPANK polypeptide in a biological sample
XX
SQ   Sequence 991 AA;
```

```
  Query Match             81.5%;  Score 4703.5;  DB 4;  Length 991;
  Best Local Similarity   82.4%;  Pred. No. 0;
  Matches  906;  Conservative   10;  Mismatches    9;  Indels  175;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||:|||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLQHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGDYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLHHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRIEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              |||||||||||| |||||:|||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTAQLKERLSYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              |||||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANTNEKTKEFLTPLHVASENAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
```

```
FT   Misc-difference 236. .237
FT                   /label= XaaXaa
FT                   /note= "encoded by GANNNT"
FT   Misc-difference 666
FT                   /label= Xaa
FT                   /note= "encoded by TNT"
FT   Misc-difference 724
FT                   /label= Xaa
FT                   /note= "encoded by AAN"
FT   Misc-difference 726
FT                   /label= Xaa
FT                   /note= "encoded by CNT"
FT   Misc-difference 734
FT                   /label= Xaa
FT                   /note= "encoded by GAN"
FT   Misc-difference 740
FT                   /label= Xaa
FT                   /note= "encoded by NTG"
FT   Misc-difference 762. .763
FT                   /label= XaaXaa
FT                   /note= "encoded by TTNTNT"
FT   Domain          947. .993
FT                   /note= "SAM domain"
FT   Misc-difference 992
FT                   /label= Xaa
FT                   /note= "encoded by NAG"
FT   Misc-difference 1267
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1274
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1277
FT                   /label= Xaa
FT                   /note= "encoded by TGA"
FT   Misc-difference 1279
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1294
FT                   /label= Xaa
FT                   /note= "encoded by TGA"
FT   Misc-difference 1314
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1322. .1323
FT                   /label= XaaXaa
FT                   /note= "encoded by TAATAA"
FT   Misc-difference 1353
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1358
FT                   /label= Xaa
FT                   /note= "encoded by TGA"
FT   Misc-difference 1371
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1377
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1380
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1393
FT                   /label= Xaa
FT                   /note= "encoded by TCA"
FT   Misc-difference 1399
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1403
FT                   /label= Xaa
FT                   /note= "encoded by TAA"
FT   Misc-difference 1408
FT                   /label= Xaa
FT                   /note= "encoded by TAG"
FT   Misc-difference 1419
FT                   /label= Xaa
FT                   /note= "encoded by TAA; Xaa=unknown"
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   N-PSDB; AAC66824.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Disclosure; Fig 3; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II protein
CC   sequence. The protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 1431 AA;

  Query Match           84.2%;  Score 4859;  DB 3;  Length 1431;
  Best Local Similarity   86.3%;  Pred. No. 0;
  Matches  951;  Conservative    6;  Mismatches  141;  Indels    4;  Gaps    2;

Qy          2 FGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 61
              |||||:  ||| |||:|| || ||||||||||||| |  : ||| | | ||||||||||
Db        166 FGRKDLXXYLLXNGANVQXRDXGGLIPLHNACSFGXAXXIXLLLXHXAXPNARDNWNYTP 225

Qy         62 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 121
                |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 XXEAAIKGKIXXCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 285

Qy        122 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 345

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 405

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 465

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 525

Qy        362 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 585

Qy        422 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 481
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 645
```

```
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838

Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898

Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958

Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018

Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078

Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095

Db       1298 EQAYPEYLITYQIMKPE 1314
```

```
RESULT 24
AAB27210
ID   AAB27210 standard; protein; 1431 AA.
XX
AC   AAB27210;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II protein sequence SEQ ID NO: 4.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 41
FT                   /label= Xaa
FT                   /note= "encoded by ANG"
FT   Misc-difference 97
FT                   /label= Xaa
FT                   /note= "encoded by ANG"
FT   Misc-difference 109
FT                   /label= Xaa
FT                   /note= "encoded by NCG"
FT   Misc-difference 113
FT                   /label= Xaa
FT                   /note= "encoded by ANC"
FT   Domain          122. .935
FT                   /note= "ANK domain"
FT   Misc-difference 125
FT                   /label= Xaa
FT                   /note= "encoded by AAN"
FT   Misc-difference 138
FT                   /label= Xaa
FT                   /note= "encoded by NTC"
FT   Misc-difference 165
FT                   /label= Xaa
FT                   /note= "encoded by NGT"
FT   Misc-difference 172. .173
FT                   /label= XaaXaa
FT                   /note= "encoded by NTTRAA"
FT   Misc-difference 177
FT                   /label= Xaa
FT                   /note= "encoded by CAN"
FT   Misc-difference 184
FT                   /label= Xaa
FT                   /note= "encoded by NCA"
FT   Misc-difference 187
FT                   /label= Xaa
FT                   /note= "encoded by NAT"
FT   Misc-difference 201
FT                   /label= Xaa
FT                   /note= "encoded by CMT"
FT   Misc-difference 203. .204
FT                   /label= XaaXaa
FT                   /note= "encoded by RAAANT"
FT   Misc-difference 206
FT                   /label= Xaa
FT                   /note= "encoded by NAT"
FT   Misc-difference 210
FT                   /label= Xaa
FT                   /note= "encoded by CNA"
FT   Misc-difference 212
FT                   /label= Xaa
FT                   /note= "encoded by NGT"
FT   Misc-difference 214
FT                   /label= Xaa
FT                   /note= "encoded by NAM"
FT   Misc-difference 226. .227
FT                   /label= XaaXaa
FT                   /note= "encoded by CNCNAT"
```

```
Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||:  || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||||| ||:||||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:||||:||| ||||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:|||||||||||||||||||:|||||||||:|||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:|||||||:|||||||||||||||||||:|||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||| ||||||||||||||||||||||||||||||||:|||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: |||  |||| |||:|:|||           : :| |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              :||:  :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:|||||||||||||||||||: |||| ||||| :    ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :|||||||||:|||||||:||||||||:::|||| |||| ||: ||:||||||||||
Db       1118 YQSVEEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              |||||||||||:|||||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              |||||||||:|||||||||||||| |||||:||||||| |||||||| ||||||||
Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095
              |||||||||||||:||
Db       1298 EQAYPEYLITYQIMKPE 1314
```

RESULT 23
AAU79537
ID AAU79537 standard; protein; 1327 AA.
XX
AC AAU79537;
XX
DT 12-AUG-2002 (first entry)
XX
DE Human tankyrase.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; cytostatic; ADP ribosylation;
KW telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia; enzyme.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Domain   1. .180
FT   /label= HPS_domain
FT   /note= "domain rich in Pro, His and Ser"
FT Domain   181. .1010
FT   /label= ANK_repeat_domain
FT Region   336. .1163
FT   /label= TRIL4
FT   /note= "Recombinant protein used for antibody production"
FT Region   436. .797
FT   /label= TRIL12
FT   /note= "Recombinant protein used for antibody production"
FT Region   973. .1149
FT   /label= ANK2
FT   /note= "Recombinant protein used for antibody production"
FT Region   1023. .1088
FT   /label= SAM_motif
FT Domain   1176. .1314
FT   /label= PARP-related_domain
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR N-PSDB; ABK48629.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Claim 3; Fig 1B; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC sequence presented is the human tankyrase
XX
SQ Sequence 1327 AA;

Query Match 84.8%; Score 4889.5; DB 5; Length 1327;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||:|||  ||||||||||||
```

```
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838

Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898

Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958

Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018

Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078

Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095

Db       1298 EQAYPEYLITYQIMKPE 1314
```

RESULT 22
AAB66279
ID AAB66279 standard; protein; 1327 AA.
XX
AC AAB66279;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase1 SEQ ID NO: 4.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63838.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 118-121; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1327 AA;

Query Match 84.8%; Score 4889.5; DB 4; Length 1327;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 910; Conservative 98; Mismatches 80; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524
```

```
CC   factor-1) binding protein. This was isolated using a two-hybrid screen
CC   with TRF1. Tankyrase is involved in regulation of telomere length through
CC   TRF1 and may directly modulate the effect of TRF1. Recombinant tankyrase,
CC   or its fragments are used to identify specific modulators which are
CC   potential drugs for countering telomere shortening associated with aging
CC   (e.g. atrophy of skin, macular degeneration or atherosclerosis) or
CC   diseases like ataxia telangiectasia. They are also used to raise specific
CC   antibodies (used as immunoassay reagents and as modulators of tankyrase
CC   activity) and in rational drug design
XX
SQ   Sequence 1327 AA;

  Query Match             84.8%;  Score 4889.5;  DB 3;  Length 1327;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  910;  Conservative   98;  Mismatches   80;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||||:|||  ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||:| ||||||::||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||||:||||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:|||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||:|: |:|||||:||| ||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||:|||||||||||||||||||:|||||||||:|||||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:||||||||:|||||||||||||||||||:|||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||| ||||||||||||||||||||||||||||||||||||:|||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: |||  |||| |||:|:|||          : :| |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              :||:  :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:|||||||||||||||||||: |||| ||||| :    ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :|||||||||:|||||||:|||||||:::|||| |||| ||: ||:|||||||||||
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              |||||||||:|||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              ||||||||||:|||||||||||||| |||||:||||||| |||||||| ||||||||
Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095
              |||||||||||||:||
Db       1298 EQAYPEYLITYQIMKPE 1314


RESULT 21
AAB27212
ID   AAB27212 standard; protein; 1327 AA.
XX
AC   AAB27212;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase I protein sequence SEQ ID NO: 8.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Claim 4; Fig 5; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase I protein
CC   sequence. The invention relates to the isoaltion of the protein and
CC   coding sequences of human tankyrase II. This protein is thought to be
CC   involved in signal transduction in the cell, and to have binding activity
CC   for other telomere-associated proteins. It is possible that it plays a
CC   role in the regulation of telomere length, thus affecting the replicative
CC   ability of the cell. The protein is useful for ribosylating target
CC   proteins, for determining tankyrase II binding activity in a sample, and
CC   for modulating telomere length in a cell
XX
SQ   Sequence 1327 AA;

  Query Match             84.8%;  Score 4889.5;  DB 3;  Length 1327;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  910;  Conservative   98;  Mismatches   80;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||||:|||  ||||||||||||
```

```
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Disclosure; Fig 2; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II protein
CC   sequence. The protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 1333 AA;

  Query Match             85.5%;  Score 4932.5;  DB 3;  Length 1333;
  Best Local Similarity   86.8%;  Pred. No. 0;
  Matches  963;  Conservative    8;  Mismatches  129;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||:: :|||||||: |   :|||||| ||||||| :|||||||||||||||||||
Db        225 GFGRKNLXKYLLQNGANXQXLYNGGLIPLHXACSFGHAKXINLLLRHGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        705 QHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||||||||||||||| ||||||| ||||| ||||| ||||||||||||||||
Db        765 ICKLLLQHGADPTKKNRDGNTLLDLVKDGXTDIQDXLRGDAVLLDAAKKGCLARVKKXXF 824

Qy        601 PDNVNCRDTQGRHSTPLHLA--AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658
              ||||||||||||||||||
Db        825 PDNVNCRDTQGRHSTPLHLAGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 884

Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        885 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 944

Qy        719 VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNL 778
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        945 XXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNL 1004

Qy        779 SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1005 SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVL 1064

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1065 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 1124

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK-------KLWERYTHRRKEVS 951
              |||||||||||||||||||||||||||||||||||||||||       : ||||||||||
Db       1125 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNXXXXXAKIRHEERYTHRRKEVS 1184

Qy        952 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1011
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1185 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1244

Qy       1012 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1071
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1245 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1304

Qy       1072 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              |||||||||||||||||||||||||||||
Db       1305 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1333


RESULT 20
AAY44402
ID   AAY44402 standard; protein; 1327 AA.
XX
AC   AAY44402;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human tankyrase.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   two-hybrid screen; telomere length regulation; recombinant tankyrase;
KW   aging; skin atrophy; macular degeneration; atherosclerosis;
KW   ataxia telangiectasia; tankyrase modulator; rational drug design.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          181. .1010
FT                   /note= "Ankyrin-specific (ANK) repeat consensus domain"
FT   Domain          1023. .1088
FT                   /label= Sterile_alpha_motif
FT   Domain          1176. .1314
FT                   /note= "Poly(ADP-ribose)polymerase (PARP) related domain"
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   N-PSDB; AAZ29627.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Claim 1; Fig 1B; 133pp; English.
XX
CC   The present sequence is human tankyrase, a TRF1 (telomeric repeat binding
```

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        129 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 188

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 248

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 308

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 368

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 428

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 488

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 548

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 608

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 668

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 728

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 788

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 848

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 908

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 968

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1028

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENH 955
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENH 1083
```

```
RESULT 19
AAB27209
ID   AAB27209 standard; protein; 1333 AA.
XX
AC   AAB27209;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II SEQ ID NO: 2.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .180
FT                   /note= "ANK domain"
FT   Domain          181. .1017
FT                   /note= "ANK domain"
FT   Misc-difference 199
FT                   /label= Xaa
FT                   /note= "encoded by CNN"
FT   Misc-difference 208. .209
FT                   /label= XaaXaa
FT                   /note= "encoded by ANCCNC"
FT   Misc-difference 212
FT                   /label= Xaa
FT                   /note= "encoded by GNG"
FT   Misc-difference 232
FT                   /label= Xaa
FT                   /note= "encoded by NTT"
FT   Misc-difference 242
FT                   /label= Xaa
FT                   /note= "encoded by NTC"
FT   Misc-difference 244
FT                   /label= Xaa
FT                   /note= "encoded by NCA"
FT   Misc-difference 255
FT                   /label= Xaa
FT                   /note= "encoded by ANT"
FT   Misc-difference 264
FT                   /label= Xaa
FT                   /note= "encoded by ANT"
FT   Misc-difference 726
FT                   /label= Xaa
FT                   /note= "encoded by TNT"
FT   Misc-difference 794
FT                   /label= Xaa
FT                   /note= "encoded by GAN"
FT   Misc-difference 800
FT                   /label= Xaa
FT                   /note= "encoded by NTG"
FT   Misc-difference 822. .823
FT                   /label= XaaXaa
FT                   /note= "encoded by TTNTNT"
FT   Domain          1030. .1095
FT                   /note= "SAM domain"
FT   Misc-difference 1055
FT                   /label= Xaa
FT                   /note= "encoded by NAG; Xaa=unknown"
FT   Domain          1190. .1328
FT                   /note= "PARP domain"
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   N-PSDB; AAC66823.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
```

```
XX
SQ   Sequence 1166 AA;

  Query Match             90.2%;  Score 5201;  DB 4;  Length 1166;
  Best Local Similarity   87.3%;  Pred. No. 0;
  Matches 1025;  Conservative    1;  Mismatches    0;  Indels  148;  Gaps   10;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVL-------------- 106
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLKMMALLTPLNVNCH 186

Qy        107 ----TGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQ 162
                  ||||||||||||||||||                ||||||||||||||||||||
Db        187 ASDGTGEYKKDELLESARSGNEE-----------------RKSTPLHLAAGYNRVKIVQ 228

Qy        163 LLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKN 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 LLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKN 288

Qy        223 RVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERL----------------AYEFKG 266
              ||||||||||||||||||||||||||||||||||||||                ||||||
Db        289 RVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLIKKHLSLEMVNFKHPQAYEFKG 348

Qy        267 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 326
              |||||||||||||                ||||||||||||   ||||||||||||||
Db        349 HSLLQAAREADVTR----------------THETALHCAAAS---KRKQICELLLRKGAN 389

Qy        327 INEKTK---EFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTC 383
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 INEKTKPYPEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTC 449

Qy        384 RLLLSYGCDPNIISLQGFTALQMGNENVQQLL------------------QEGISLGNSE 425
              ||||||||||||||||||||||||||||||                  ||||||||||
Db        450 RLLLSYGCDPNIISLQGFTALQMGNENVQQLLTVKKLCTVQSVNCRDIEGQEGISLGNSE 509

Qy        426 ADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGAD 485
              |||||||||||||                ||||||||||||||||||||||||||||||
Db        510 ADRQLLEAAKAGDVE-----------------RQSTPLHFAAGYNRVSVVEYLLQHGAD 551

Qy        486 VHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLL 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 VHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLL 611

Qy        546 LQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVN 605
              ||||||||||||||||||||||||||||||                ||||||||||||
Db        612 LQHGADPTKKNRDGNTPLDLVKDGDTDIQ------------------RVKKLSSPDNVN 652

Qy        606 CRDTQG------------------RHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGG 646
              ||||||                  ||||||||||||||||||||||||||||||||||||
Db        653 CRDTQGDLLRGDAALLDAAKKGCLARHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGG 712

Qy        647 LIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPT 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 LIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPT 772

Qy        707 LKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSP 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 LKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSP 832

Qy        767 SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMD 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMD 892

Qy        827 IFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSG 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 IFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSG 952

Qy        887 TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHR 946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        953 TILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHR 1012

Qy        947 RKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYV 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 RKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYV 1072

Qy       1007 YGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVN 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 YGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVN 1132

Qy       1067 GLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||||||||||||||
Db       1133 GLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1166


RESULT 18
AAE35352
ID   AAE35352 standard; protein; 1083 AA.
XX
AC   AAE35352;
XX
DT   17-JUN-2003  (first entry)
XX
DE   Human colon cancer-associated polypeptide, TNKL (BC-203).
XX
KW   Human; colon cancer-associated polypeptide; immune response; therapy;
KW   colon cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 18
FT                   /note= "Encoded by ATCC"
FT   Misc-difference 77
FT                   /note= "Encoded by GAGC"
XX
PN   WO200290986-A1.
XX
PD   14-NOV-2002.
XX
PF   02-MAY-2002; 2002WO-US013994.
XX
PR   04-MAY-2001; 2001US-00849602.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chen Y,  Old LJ,  Scanlan MJ,  Stockert E;
XX
DR   WPI; 2003-112003/10.
DR   N-PSDB; AAD54031.
XX
PT   Diagnosing colon cancer in a subject comprises identifying colon cancer-
PT   associated polypeptides as antigens that elicit immune response in colon
PT   cancer.
XX
PS   Example 1; Page 118-121; 122pp; English.
XX
CC   The invention relates to a method for diagnosing colon cancer in a
CC   subject which comprises identifying colon cancer-associated polypeptides
CC   as antigens that elicit immune response in colon cancer. The method is
CC   useful for diagnosing, determining onset, progression, or regression of
CC   colon cancer in a subject, or for selecting a course of treatment of a
CC   subject having or suspected of having colon cancer. The colon cancer-
CC   associated polypeptides are useful as markers for diagnosing colon
CC   cancer, and for following the course of treatment of colon cancer. The
CC   present sequence is human colon cancer-associated polypeptide
XX
SQ   Sequence 1083 AA;

  Query Match             86.2%;  Score 4974;  DB 6;  Length 1083;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  954;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      27 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 IPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTI 60

Qy      87 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 RNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKS 120

Qy     147 TPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNA 206
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||| ||||
Db     121 TPLHLAAGYNRVKIVQLLLQHGRDVHAKDKGDLVPLHNACSYGHYEVTELLVKHGGCVNA 180

Qy     207 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKG 266
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db     181 MDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCKNKSAIDLAPTPQLKERLAYEFKG 240

Qy     267 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 326
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 HSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN 300

Qy     327 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 360

Qy     387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 420

Qy     447 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 506
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHY 480

Qy     507 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 566
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 EVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLV 540

Qy     567 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 KDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNL 600

Qy     627 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEA 686
           |||||||||||||||||||||||||||||||||||||||| :||||||||||||||||
Db     601 EVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNASLNATDKWAFTPLHEA 660

Qy     687 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVL 720

Qy     747 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 NGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVP 780

Qy     807 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 866
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVER 840

Qy     867 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 926
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 LISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY 900

Qy     927 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 986
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIG 960

Qy     987 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1046
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSA 1020

Qy    1047 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 MKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1074
```

```
RESULT 17
AAB47022
ID   AAB47022 standard; protein; 1166 AA.
XX
AC   AAB47022;
XX
DT   29-MAR-2001  (first entry)
XX
DE   Human SPANK.
XX
KW   SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW   poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW   ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW   adipocyte; insulin signalling pathway; hyperlipidaemia;
KW   glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW   cardiac insufficiency; coronary insufficiency; stroke;
KW   high blood pressure; non-insulin dependent diabetes; hypertension;
KW   hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 173. .209
FT                   /note= "Encoded by nucleotides 768. .878"
FT   Misc-difference 327. .362
FT                   /note= "Encoded by nucleotides 1230. .1337"
FT   Misc-difference 375. .398
FT                   /note= "Encoded by nucleotides 1374. .1445"
FT   Misc-difference 482. .524
FT                   /note= "Encoded by nucleotides 1695. .1823"
FT   Misc-difference 641. .677
FT                   /note= "Encoded by nucleotides 2172. .2282"
XX
PN   WO200077225-A1.
XX
PD   21-DEC-2000.
XX
PF   09-JUN-2000; 2000WO-US015926.
XX
PR   11-JUN-1999;   99US-0138957P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Chi N,  Lodish HF;
XX
DR   WPI; 2001-091404/10.
DR   N-PSDB; AAC85294.
XX
PT   New insulin signaling protein SPANK, useful for reducing body mass,
PT   glucose intolerance or insulin resistance and for preventing or treating
PT   obesity-related and muscle-related diseases.
XX
PS   Claim 3; Fig 3; 65pp; English.
XX
CC   This sequence represents human SPANK. The SPANK protein comprises 3
CC   domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP (poly
CC   adenosine diphosphate-ribose polymerase) catalytic domain; and (c) an ANK
CC   domain composed of ankyrin repeats. SPANK is a cytosolic protein which
CC   can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC   aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC   perinuclear region of adipocytes. It is an effector in the insulin
CC   signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC   mass, reducing glucose intolerance or insulin resistance, for preventing
CC   or treating obesity-related diseases or disorders, such as obesity,
CC   cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC   atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC   dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC   also useful for preventing or treating muscle-related diseases or
CC   disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC   Antibodies immunospecific for SPANK are useful for detecting the presence
CC   of SPANK polypeptide in a biological sample
```

```
CC   sequence represents the tankyrase H isoform 1, Taho E/A dC, which
CC   contains a Glu to Ala mutation and has a truncated C-terminal containing
CC   a 2 amino acid insertion described in the disclosure of the invention
XX
SQ   Sequence 1084 AA;

  Query Match             98.1%;  Score 5661;  DB 6;  Length 1084;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1080;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAAYVIYRGEQ 1080

Qy       1081 AYPE 1084
              |  :
Db       1081 ALSZ 1084


RESULT 16
AAY05734
ID   AAY05734 standard; protein; 1074 AA.
XX
AC   AAY05734;
XX
DT   19-JUL-1999  (first entry)
XX
DE   Human Grb7 effector 2.2412 protein.
XX
KW   Grb7 effector; 2.2412 protein; human; signal transduction; tumour marker;
KW   breast cancer; prostate cancer; prognosis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO9915647-A1.
XX
PD   01-APR-1999.
XX
PF   23-SEP-1998;   98WO-AU000795.
XX
PR   23-SEP-1997;   97AU-00009388.
XX
PA   (GARV-) GARVAN INST MEDICAL RES.
XX
PI   Daly RJ,  Sutherland RL;
XX
DR   WPI; 1999-254707/21.
DR   N-PSDB; AAX25366.
XX
PT   New candidate effector for the Grb7 family of signaling proteins, and
PT   specific antibody, useful for detection and treatment of cancer.
XX
PS   Claim 9; Fig 1; 24pp; English.
XX
CC   The present sequence represents a novel candidate effector for the Grb7
CC   family of signalling proteins, termed 2.2412. The sequence is predicted
CC   from a partial cDNA (see AAX25366). 2.2412 mRNA transcripts (7 kb) were
CC   detected in all tissues examined with the exception of kidney. Expression
CC   was particularly high in skeletal muscle and placenta. Analysis of the
CC   sequence revealed significant homology to a large number of proteins
CC   containing ankyrin-like repeats. The 2.2412 gene was localised to between
CC   chromosome 10q23.2 and proximal 10q23.32. Deletions in the 10q22-25
CC   region have been detected in human breast, prostate, renal, small cell
CC   lung and endometrial carcinomas, glioblastoma multiforme, melanoma and
CC   meningiomas. Detection of the protein encoded by the 2.2412 cDNA in a
CC   sample should provide a useful tumour marker and/or prognostic indicator
CC   for certain human cancers, in particular breast cancer and prostate
CC   cancer. Antagonism of the interaction between Grb7 family members and the
CC   encoded protein should provide a novel treatment stategy for human
CC   diseases exhibiting aberrant receptor tyrosine kinase signalling, such as
CC   cancer. Anti-Grb7 antibodies can be used in methods of detecting the
CC   presence of 2.2412 protein in a sample
XX
SQ   Sequence 1074 AA;

  Query Match             97.2%;  Score 5605;  DB 2;  Length 1074;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1069;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKSCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALXLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRXLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166
```

RESULT 15
ABG73736
ID ABG73736 standard; protein; 1084 AA.
XX
AC ABG73736;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 1 TH-1 variant Taho E/A dC.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TH-1; TaHo;
KW cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .22
FT		/label= Ankyrin_repeat_1
FT	Region	24. .55
FT		/label= Ankyrin_repeat_2
FT	Region	57. .87
FT		/label= Ankyrin_repeat_3
FT	Region	89. .113
FT		/label= Ankyrin_repeat_4
FT	Region	115. .139
FT		/label= Ankyrin_repeat_5
FT	Misc-difference	127
FT		/note= "TH1 start"
FT	Region	147. .176
FT		/label= Ankyrin_repeat_6
FT	Region	178. .208
FT		/label= Ankyrin_repeat_7
FT	Region	210. .241
FT		/label= Ankyrin_repeat_8
FT	Region	243. .281
FT		/label= Ankyrin_repeat_9
FT	Region	283. .332
FT		/label= Ankyrin_repeat_10
FT	Region	334. .365
FT		/label= Ankyrin_repeat_11
FT	Region	367. .398
FT		/label= Ankyrin_repeat_12
FT	Region	400. .424
FT		/label= Ankyrin_repeat_13
FT	Region	426. .458
FT		/label= Ankyrin_repeat_14
FT	Region	460. .491
FT		/label= Ankyrin_repeat_15
FT	Region	493. .524
FT		/label= Ankyrin_repeat_16
FT	Region	526. .557
FT		/label= Ankyrin_repeat_17
FT	Region	559. .594
FT		/label= Ankyrin_repeat_18
FT	Region	596. .644
FT		/label= Ankyrin_repeat_19
FT	Region	646. .677
FT		/label= Ankyrin_repeat_20
FT	Region	679. .710
FT		/label= Ankyrin_repeat_21
FT	Region	712. .755
FT		/label= Ankyrin_repeat_22
FT	Region	757. .793
FT		/label= Ankyrin_repeat_23
FT	Domain	797. .845
FT		/label= SAM_domain
FT	Region	855. .869
FT		/note= "Unidentified region"
FT	Domain	956. .1084
FT		/label= PARP_domain
FT	Misc-difference	1072
FT		/note= "This residue is mutated from Glu (wild type, see ABG73732) to Ala"

XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Disclosure; Fig 16A-C; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
                 ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166
```

```
RESULT 14
AAY72589
ID   AAY72589 standard; protein; 1166 AA.
XX
AC   AAY72589;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human tankyrase homolog protein (THP).
XX
KW   Human; tankyrase homolog protein; THP; gene therapy; cancer; tumour;
KW   basal cell carcinoma; therapy; genetic mapping; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 368
FT                   /label= Met, Leu, Val
FT   Misc-difference 392
FT                   /label= Asn, Thr
FT   Misc-difference 415
FT                   /label= Val, Ile
FT   Misc-difference 558
FT                   /label= Gly, Glu
FT   Misc-difference 559
FT                   /label= Gly, Ala
FT   Misc-difference 764
FT                   /label= Leu, Val
FT   Misc-difference 884
FT                   /label= Asn, His, Asp, Tyr
XX
PN   WO200104326-A1.
XX
PD   18-JAN-2001.
XX
PF   03-JUL-2000; 2000WO-EP006609.
XX
PR   09-JUL-1999;   99US-00350982.
XX
PA   (PHAA ) PHARMACIA & UPJOHN SPA.
XX
PI   Berthelsen J,  Toma S,  Isacchi A;
XX
DR   WPI; 2001-168422/17.
DR   N-PSDB; AAD02578.
XX
PT   New tankyrase homolog protein (THP) polynucleotide and polypeptide useful
PT   in gene therapy, diagnosis and treatment or prevention of unregulated
PT   cell growth, such as cancer or tumor cell growth.
XX
PS   Claim 20; Page 50-54; 60pp; English.
XX
CC   The present sequence is human tankyrase homolog protein (THP). The THP
CC   polypeptides and polynucleotides of the invention are useful in gene
CC   therapy and for treating or preventing unregulated cell growth such as
CC   cancer or tumour (e.g. basal cell carcinoma). The nucleic acid molecules
CC   of the invention and their fragments are useful for restriction fragment
CC   length polymorphism (RFLP) associated with certain disorders, as well as
CC   for genetic mapping. Antisense oligonucleotides, or fragments of nucleic
CC   acid encoding THP are useful as diagnostic tools for probing the
CC   expression of Thp gene in various tissues. THP can be used as antigens
CC   for raising antibodies against them and in assays for identifying
CC   compounds that modulate their activity. They are used in the manufacture
CC   of a medicament directed towards cancers or tumours. THP are also useful
CC   for screening compounds in a variety of drugs screening techniques and as
CC   a research tool for identification, characterisation and purification of
CC   interacting, regulatory proteins
XX
SQ   Sequence 1166 AA;
```

Query Match 99.1%; Score 5717; DB 4; Length 1166;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1091; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
                 ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
                 | ||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||
Db        367 AXHCAAASPYPKRKQICELLLRKGAXINEKTKEFLTPLHVASEKAHNDXVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
                 |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKXXLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606
```

```
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYLAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100
```

```
RESULT 13
AAB27211
ID   AAB27211 standard; protein; 1166 AA.
XX
AC   AAB27211;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II protein sequence SEQ ID NO: 6.
XX
KW   Human; tankyrase II; telomere length; signal transduction.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 124
FT                   /note= "encoded by TTA"
FT   Misc-difference 125
FT                   /note= "encoded by TAC"
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   N-PSDB; AAC66825.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Claim 4; Fig 4; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II protein
CC   sequence. The protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 1166 AA;
```

```
  Query Match             99.7%;  Score 5750;  DB 3;  Length 1166;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 1097;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNTY 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
```

```
Db       1366 AYPEYLITYQIMRPEGMVDG 1385


RESULT 12
ABG73735
ID   ABG73735 standard; protein; 1100 AA.
XX
AC   ABG73735;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TH-1 variant Taho F/L.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; mutant;
KW   TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP; variant.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .22
FT                   /label= Ankyrin_repeat_1
FT   Region          24. .55
FT                   /label= Ankyrin_repeat_2
FT   Region          57. .87
FT                   /label= Ankyrin_repeat_3
FT   Region          89. .113
FT                   /label= Ankyrin_repeat_4
FT   Region          115. .139
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 127
FT                   /note= "TH1 start"
FT   Region          147. .176
FT                   /label= Ankyrin_repeat_6
FT   Region          178. .208
FT                   /label= Ankyrin_repeat_7
FT   Region          210. .241
FT                   /label= Ankyrin_repeat_8
FT   Region          243. .281
FT                   /label= Ankyrin_repeat_9
FT   Region          283. .332
FT                   /label= Ankyrin_repeat_10
FT   Region          334. .365
FT                   /label= Ankyrin_repeat_11
FT   Region          367. .398
FT                   /label= Ankyrin_repeat_12
FT   Region          400. .424
FT                   /label= Ankyrin_repeat_13
FT   Region          426. .458
FT                   /label= Ankyrin_repeat_14
FT   Region          460. .491
FT                   /label= Ankyrin_repeat_15
FT   Region          493. .524
FT                   /label= Ankyrin_repeat_16
FT   Region          526. .557
FT                   /label= Ankyrin_repeat_17
FT   Region          559. .594
FT                   /label= Ankyrin_repeat_18
FT   Region          596. .644
FT                   /label= Ankyrin_repeat_19
FT   Region          646. .677
FT                   /label= Ankyrin_repeat_20
FT   Region          679. .710
FT                   /label= Ankyrin_repeat_21
FT   Region          712. .755
FT                   /label= Ankyrin_repeat_22
FT   Region          757. .793
FT                   /label= Ankyrin_repeat_23
FT   Domain          797. .845
FT                   /label= SAM_domain
FT   Region          855. .869
FT                   /note= "Unidentified region"
FT   Domain          956. .1094
FT                   /label= PARP_domain
FT   Misc-difference 995
FT                   /note= "This residue is mutated from Phe (wild type, see
FT                   ABG73732) to Leu"
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-C; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform 1, TH-1 variant Taho F/L
CC   described in the disclosure of the invention
XX
SQ   Sequence 1100 AA;

  Query Match            99.9%;  Score 5763;  DB 6;  Length 1100;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
```

```
Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1062

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1122

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1182

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1242

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1243 AYPEYLITYQIMRPEGMVDG 1262
```

RESULT 11
AAB66294
ID AAB66294 standard; protein; 1385 AA.
XX
AC AAB66294;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-LONG SEQ ID NO: 133.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63952.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Claim 2; Page 191-194; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1385 AA;

```
  Query Match             99.9%;  Score 5766;  DB 4;  Length 1385;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        286 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 345

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 405

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 465

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 525

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 585

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 645

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 705

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 765

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 825

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 885

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        886 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 945

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        946 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 1005

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1006 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 1065

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 1125

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1185

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1245

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1305

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1365

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
```

```
Db        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1221 AYPEYLITYQIMRPEGMVDG 1240
```

RESULT 10
AAB66290
ID AAB66290 standard; protein; 1262 AA.
XX
AC AAB66290;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus protein SEQ ID NO: 107.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63930.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 2; Page 173-176; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1262 AA;

Query Match 99.9%; Score 5766; DB 4; Length 1262;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        163 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 222

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 282

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 342

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 402

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 462

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 522

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 582

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 642

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 702

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 762

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 822

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 882

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 942

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 1002
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
                 ||||||||||||||||||||
Db       1221 AYPEYLITYQIMRPEGMVDG 1240
```

RESULT 9
ABG73731

```
ID   ABG73731 standard; protein; 1240 AA.
XX
AC   ABG73731;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 2 TaHo-2 fragment from clone K23.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-2;
KW   TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS   Unidentified.
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y,  Chan E,  Xu X,  Huang B,  Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
DR   N-PSDB; ABQ77067.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Claim 14; Fig 4; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents the tankyrase H isoform 2, TaHo-2, isolated from
CC   clone K23 which is described in the disclosure of the invention
XX
SQ   Sequence 1240 AA;
```

Query Match 99.9%; Score 5766; DB 6; Length 1240;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
                 ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 369

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 429

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 489

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 549

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 609

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 669

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 729

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 789

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 849

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 909

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 969

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1029

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1089

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1149

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1150 AYPEYLITYQIMRPEGMVDG 1169
```

RESULT 8
AAY97749
ID AAY97749 standard; protein; 1240 AA.
XX
AC AAY97749;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 2 protein sequence.
XX
KW Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR N-PSDB; AAA91488.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 22; Fig 4; 63pp; English.
XX
CC This sequence is the Tankyrase homologue isotype 2 (TaHo-2) protein of
CC the invention. The invention also relates to the TaHo-2 protein. The TaHo
CC proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as
CC vaccine, and for construction of hybridisation probes for mapping the
CC gene which encodes TaHo and for the genetic analysis of individuals with
CC genetic disorders. The TaHo proteins, and their coding sequences are
CC useful in screening assays
XX
SQ Sequence 1240 AA;

Query Match 99.9%; Score 5766; DB 4; Length 1240;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
```

```
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         70 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 129

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 189

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 249

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 309

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 369

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 429

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 489

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 549

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 609

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        610 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 669

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 729

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        730 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 789

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        790 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 849

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        850 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 909

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        910 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 969

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        970 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1029

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1030 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1089

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1090 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1149

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1150 AYPEYLITYQIMRPEGMVDG 1169


RESULT 7
AAB66278
ID   AAB66278 standard; protein; 1169 AA.
XX
AC   AAB66278;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 related protein sequence SEQ ID NO: 2.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63837.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Disclosure; Page 109-113; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1169 AA;

  Query Match             99.9%;  Score 5766;  DB 4;  Length 1169;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         70 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 129

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 189

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 249

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 309
```

```
CC   ABZ12066) or their mature protein coding portion, active domain coding
CC   protein or complementary sequences. The polynucleotides are useful for
CC   identifying expressed genes or for physical mapping of human genome. The
CC   encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC   markers, as a food supplement, for generating antibodies, in medical
CC   imaging, screening and diagnostic assays and for treating cell-
CC   proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC   or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC   diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC   platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC   or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC   arthritis, etc. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1166 AA;

  Query Match             99.9%;  Score 5766;  DB 5;  Length 1166;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166


RESULT 6
AAB66288
ID   AAB66288 standard; protein; 1169 AA.
XX
AC   AAB66288;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone consensus protein SEQ ID NO: 101.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   N-PSDB; AAF63926.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 1; Page 162-1665; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1169 AA;

  Query Match             99.9%;  Score 5766;  DB 4;  Length 1169;
```

```
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Claim 3; Page 200-203; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1166 AA;

  Query Match             99.9%;  Score 5766;  DB 4;  Length 1166;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166


RESULT 5
ABP69457
ID   ABP69457 standard; protein; 1166 AA.
XX
AC   ABP69457;
XX
DT   20-JAN-2003  (first entry)
XX
DE   Human polypeptide SEQ ID NO 1504.
XX
KW   Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW   cell-proliferative disorder; neurodegenerative disease; bacterial;
KW   Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW   multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW   arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW   antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW   haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
KW   antiarthritic.
XX
OS   Homo sapiens.
XX
PN   WO200270539-A2.
XX
PD   12-SEP-2002.
XX
PF   05-MAR-2002; 2002WO-US005095.
XX
PR   05-MAR-2001; 2001US-00799451.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Zhou P,  Goodrich RW,  Asundi V,  Zhang J,  Zhao QA,  Ren F;
PI   Xue AJ,  Yang Y,  Ma Y,  Yamazaki V,  Chen R,  Wang Z,  Ghosh M;
PI   Wehrman T,  Wang J,  Wang D,  Drmanac RT;
XX
DR   WPI; 2002-759812/82.
DR   N-PSDB; ABZ11674.
XX
PT   New polynucleotides comprising sequences assembled from expressed
PT   sequence tags (ESTs), useful for treating cell-proliferative,
PT   neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT   or coagulation disorders.
XX
PS   Claim 9; SEQ ID NO 1504; 1012pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated polynucleotide (I) comprising a
CC   nucleotide sequence selected from any of 948 sequences (ABZ11119-
```

PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
DR N-PSDB; ABQ77066.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Claim 14; Fig 3; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence represents the tankyrase H isoform 1, TaHo-1, isolated from
CC clone TH-1 which is described in the disclosure of the invention
XX
SQ Sequence 1100 AA;

Query Match 100.0%; Score 5769; DB 6; Length 1100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100
```

RESULT 4
AAB66295
ID AAB66295 standard; protein; 1166 AA.
XX
AC AAB66295;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR N-PSDB; AAF63953.

```
FT                   /label= Ankyrin_repeat_20
FT   Region          679. .710
FT                   /label= Ankyrin_repeat_21
FT   Region          712. .755
FT                   /label= Ankyrin_repeat_22
FT   Region          757. .793
FT                   /label= Ankyrin_repeat_23
FT   Domain          797. .845
FT                   /label= SAM_domain
FT   Region          855. .869
FT                   /note= "Unidentified region"
FT   Domain          956. .1094
FT                   /label= PARP_domain
XX
PN   WO200286170-A1.
XX
PD   31-OCT-2002.
XX
PF   25-APR-2002; 2002WO-US013185.
XX
PR   25-APR-2001; 2001US-00843159.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR   WPI; 2003-093158/08.
XX
PT   New recombinant nucleic acid encoding a cell cycle protein, useful for
PT   diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS   Disclosure; Fig 16A-C; 90pp; English.
XX
CC   This invention describes a novel recombinant nucleic acid encoding the
CC   cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC   invention have cytostatic and immunostimulant activity and can be used
CC   for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC   nucleic acid encoding the protein are useful for diagnosing and treating
CC   a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC   tumour cell. The methods are useful for screening for a bioactive agent
CC   capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC   agent capable of modulating a cell cycle protein tankyrase H or PARP
CC   (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC   inducing an immune response against the cell cycle proteins. This
CC   sequence represents tankyrase H isoform 1, TH-1, described in the
CC   disclosure of the invention
XX
SQ   Sequence 1100 AA;

  Query Match             100.0%;  Score 5769;  DB 6;  Length 1100;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100


RESULT 3
ABG73730
ID   ABG73730 standard; protein; 1100 AA.
XX
AC   ABG73730;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TaHo-1 from clone TH-1.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-1;
KW   TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS   Unidentified.
XX
```

```
CC   vaccine, and for construction of hybridisation probes for mapping the
CC   gene which encodes TaHo and for the genetic analysis of individuals with
CC   genetic disorders. The TaHo proteins, and their coding sequences are
CC   useful in screening assays
XX
SQ   Sequence 1100 AA;

  Query Match             100.0%;  Score 5769;  DB 4;  Length 1100;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100


RESULT 2
ABG73732
ID   ABG73732 standard; protein; 1100 AA.
XX
AC   ABG73732;
XX
DT   08-APR-2003  (first entry)
XX
DE   Tankyrase H isotype 1 TH-1.
XX
KW   Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TH-1; TaHo;
KW   cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW   cell cycle related disorder; poly ADP-ribose polymerase; PARP.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .22
FT                   /label= Ankyrin_repeat_1
FT   Region          24. .55
FT                   /label= Ankyrin_repeat_2
FT   Region          57. .87
FT                   /label= Ankyrin_repeat_3
FT   Region          89. .113
FT                   /label= Ankyrin_repeat_4
FT   Region          115. .139
FT                   /label= Ankyrin_repeat_5
FT   Misc-difference 127
FT                   /note= "TH1 start"
FT   Region          147. .176
FT                   /label= Ankyrin_repeat_6
FT   Region          178. .208
FT                   /label= Ankyrin_repeat_7
FT   Region          210. .241
FT                   /label= Ankyrin_repeat_8
FT   Region          243. .281
FT                   /label= Ankyrin_repeat_9
FT   Region          283. .332
FT                   /label= Ankyrin_repeat_10
FT   Region          334. .365
FT                   /label= Ankyrin_repeat_11
FT   Region          367. .398
FT                   /label= Ankyrin_repeat_12
FT   Region          400. .424
FT                   /label= Ankyrin_repeat_13
FT   Region          426. .458
FT                   /label= Ankyrin_repeat_14
FT   Region          460. .491
FT                   /label= Ankyrin_repeat_15
FT   Region          493. .524
FT                   /label= Ankyrin_repeat_16
FT   Region          526. .557
FT                   /label= Ankyrin_repeat_17
FT   Region          559. .594
FT                   /label= Ankyrin_repeat_18
FT   Region          596. .644
FT                   /label= Ankyrin_repeat_19
FT   Region          646. .677
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:36:25 ; Search time 52.1795 Seconds
(without alignments)
5956.407 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5769	100.0	1100	4	AAY97748	Aay97748 Tankyrase
2	5769	100.0	1100	6	ABG73732	Abg73732 Tankyrase
3	5769	100.0	1100	6	ABG73730	Abg73730 Tankyrase
4	5766	99.9	1166	4	AAB66295	Aab66295 Human tan
5	5766	99.9	1166	5	ABP69457	Abp69457 Human pol
6	5766	99.9	1169	4	AAB66288	Aab66288 Human tan
7	5766	99.9	1169	4	AAB66278	Aab66278 Human tan
8	5766	99.9	1240	4	AAY97749	Aay97749 Tankyrase
9	5766	99.9	1240	6	ABG73731	Abg73731 Tankyrase
10	5766	99.9	1262	4	AAB66290	Aab66290 Human tan
11	5766	99.9	1385	4	AAB66294	Aab66294 Human tan
12	5763	99.9	1100	6	ABG73735	Abg73735 Tankyrase
13	5750	99.7	1166	3	AAB27211	Aab27211 Human tan
14	5717	99.1	1166	4	AAY72589	Aay72589 Human tan
15	5661	98.1	1084	6	ABG73736	Abg73736 Tankyrase
16	5605	97.2	1074	2	AAY05734	Aay05734 Human Grb
17	5201	90.2	1166	4	AAB47022	Aab47022 Human SPA
18	4974	86.2	1083	6	AAE35352	Aae35352 Human col
19	4932.5	85.5	1333	3	AAB27209	Aab27209 Human tan
20	4889.5	84.8	1327	3	AAY44402	Aay44402 Human tan
21	4889.5	84.8	1327	3	AAB27212	Aab27212 Human tan
22	4889.5	84.8	1327	4	AAB66279	Aab66279 Human tan
23	4889.5	84.8	1327	5	AAU79537	Aau79537 Human tan
24	4859	84.2	1431	3	AAB27210	Aab27210 Human tan
25	4703.5	81.5	991	4	AAB47023	Aab47023 Mouse SPA
26	4624.5	80.2	907	4	AAB48574	Aab48574 Human bre
27	4102	71.1	784	4	AAB66285	Aab66285 Human tan
28	3959	68.6	756	4	AAB66286	Aab66286 Human tan
29	3835.5	66.5	1181	4	ABB60894	Abb60894 Drosophil
30	3835.5	66.5	1181	4	AAB66297	Aab66297 Drosophil
31	3596	62.3	802	5	AAO20512	Aao20512 Protein o
32	3093	53.6	949	3	AAY44404	Aay44404 Human tru
33	3093	53.6	949	5	AAU79539	Aau79539 Truncated
34	2916	50.5	584	6	ABG73734	Abg73734 Tankyrase
35	2367	41.0	522	4	AAB66287	Aab66287 Human tan
36	2165	37.5	415	4	ABB11320	Abb11320 Human Grb
37	2053.5	35.6	1099	4	AAB66301	Aab66301 Human tan
38	1962	34.0	373	6	ABU70878	Abu70878 Human adi
39	1942	33.7	385	4	AAB48571	Aab48571 Human T c
40	1918.5	33.3	673	3	AAY44403	Aay44403 Human tru
41	1918.5	33.3	673	5	AAU79538	Aau79538 Truncated
42	1895	32.8	362	4	AAB66300	Aab66300 Human TRF
43	1706	29.6	331	4	AAB48572	Aab48572 Human bre
44	1467	25.4	308	6	ABU70889	Abu70889 Human adi
45	1395	24.2	261	4	AAB94719	Aab94719 Human pro

ALIGNMENTS

RESULT 1
AAY97748
ID AAY97748 standard; protein; 1100 AA.
XX
AC AAY97748;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 1 protein sequence.
XX
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine.
XX
OS Unidentified.
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR N-PSDB; AAA91487.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 22; Fig 3; 63pp; English.
XX
CC This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein of
CC the invention. The invention also relates to the TaHo-2 protein. The TaHo
CC proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as

```
Qy        643 DKGGLIPLHN---AASYGHVDVAALLIKYNACVNATDKWAFT-PLHEAAQKGRTQLCALL 698
              : || | | :   :    || |: :: : :   || | :   |  |  |  :|     ||
Db        554 ENGGKIDLRSFFESVKCGHEDITSIFLAFGFPVNQTSRDNKTSALMVAVGRGHLGTVKLL 613

Qy        699 LAHGADPTLKNQEGQTPL--------DLVSADDVSALLTAAMPPSALPS----------- 739
              |  ||||| ::::|:| |          : :::::  :  |       |
Db        614 LEAGADPTKRDKKGRTALYYAKNSIMGITNSEEIQLIENAINNYLKKHSEDNNDDDDDDD 673

Qy        740 ----CYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTE 795
                  || :           |   |: |           |:     |:    |  |: ::
Db        674 NNNETYKHEKKREKTQSPILASRRSATPRIEDEEDDTRMLNLADDDFNNDRDVKESTTSD 733

Qy        796 GASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIG-INAY 854
                |:  |             |:| :: :|  :|:   |         ::|| |  ::
Db        734 SRKRLDDNE------------NVGTQYSLDWKKRKTNAL-----QDEEKLKSISPLSME 775

Qy        855 GHRHKLIKGVE-------------RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
               |  |  | ||              ||   ::      |         :|  |: |:|  :
Db        776 PHSPKKAKSVEISKIHEETAAEREARLKEEEEYRKKRLEKKRKKEQELLQKLAEDEK--K 833

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              :||: :  | |              : :::|   :|   |   | ||: | ::  |
Db        834 RIEEQEKQKVLE--------------MERLEKATLEK--ARKMEREKEMEEISYRRAVR 876

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYV 1006
               :   |   :|      |:::  |    |    || :   |::::|
Db        877 DLYPLGLKIIN------FNDKLDY--KRFLPLYYFVD--EKNDKFV 912


RESULT 45
T50984
related to 26s proteasome subunit p28 [imported] - Neurospora crassa
N;Alternate names: protein B7F18.30
C;Species: Neurospora crassa
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 08-Sep-2002
C;Accession: T50984
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A;Reference number: Z25286
A;Accession: T50984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <SCH>
A;Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.30
A;Experimental source: BAC clone B7F18; strain OR74A
C;Genetics:
A;Gene: NCSP:B7F18.30
A;Map position: 6
A;Introns: 17/1; 25/3; 68/3
C;Superfamily: hypothetical protein YGR232w; ankyrin repeat homology

  Query Match             4.9%;  Score 284;  DB 2;  Length 237;
  Best Local Similarity   31.5%;  Pred. No. 1.5e-10;
  Matches   85;  Conservative   29;  Mismatches   92;  Indels   64;  Gaps    7;

Qy        464 LHFAAGYNRVSVVEYLLQHGADV-HAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 522
              :| ||   : |:|| ||     :   ||  | :|:| ||||   || ||||
Db          8 IHAAARDGKASIVESLLNANPKLAQRKDDDGRLPIHWACSYNRKEVVELLVN-------- 59

Qy        523 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL---VKDGDTDIQDLLRG 579
                                 | | ||  ::  | ||  :   ||| |  | |||
Db         60 ----------------------QKGFDPDVEDDMGWTPFMISASVKDSDA-IIDLL-- 92

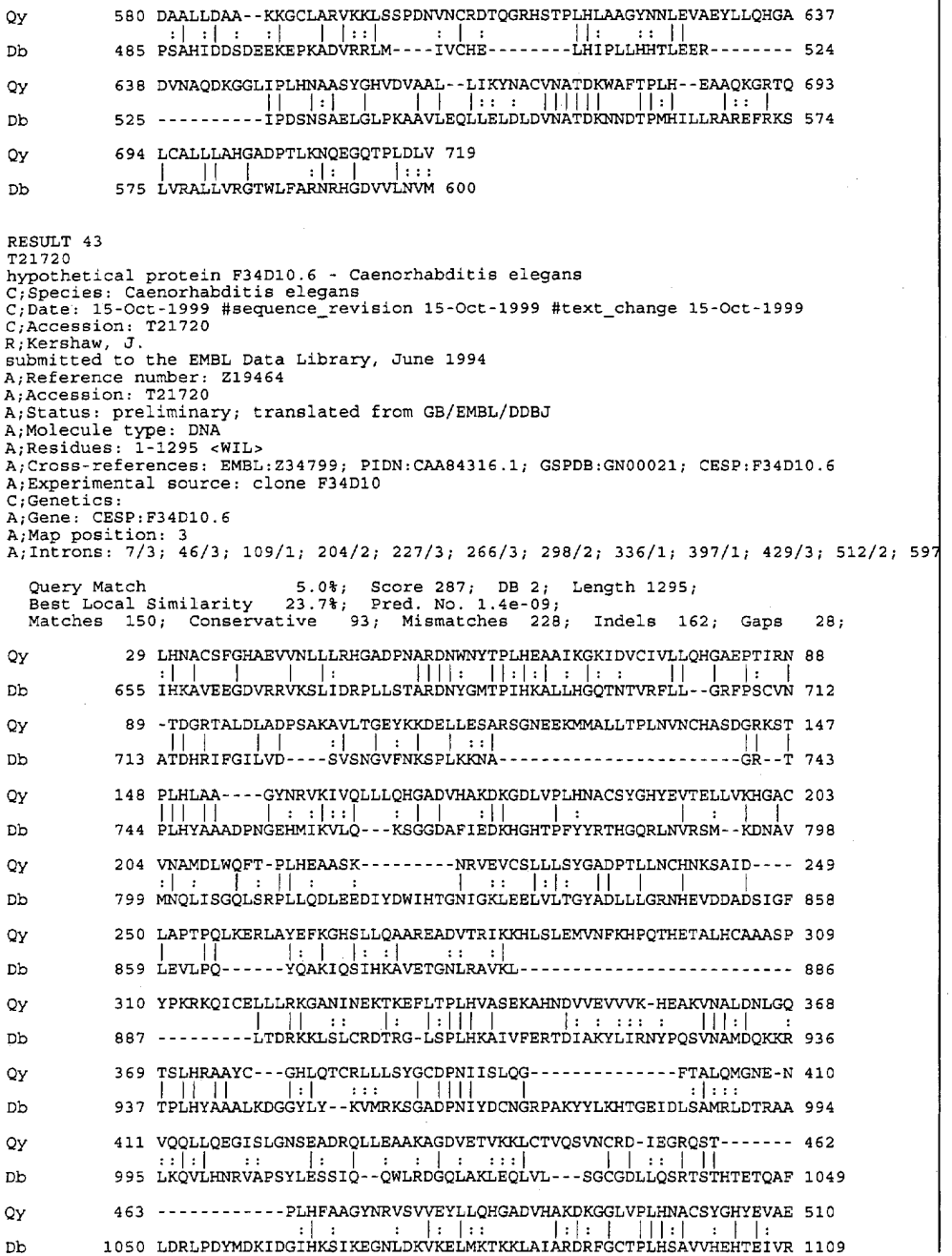
Qy        580 DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ----- 634
                           ||   ::|   |  :  | ||  |  ||:::|  ||
Db         93 -----------------LSRGADIN--QTNHQSQTALHFIASKNNIDLARKLLSPDMKP 132

Qy        635 HGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQL 694
               |  |  :|| |  ||| ||: | | :  ||::: : :||:|   :|||| |  :|
Db        133 KPASVRVKDKRGQYPLHRAAAIGSVPMINLLLQHKSPINASDNAGYTPLHHAVAEGHGHA 192

Qy        695 CALLLAHGADPTLKNQEGQTPLDLVSADDV 724
                  ||  ||:   |: :|   |||    ||

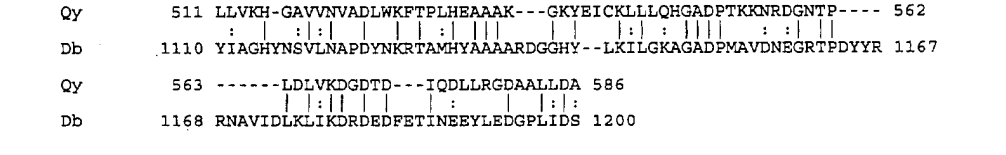
Db        193 AVALLKAGAETDKKDMDGYLALDLAPDKDV 222


Search completed: March 25, 2004, 16:44:52
Job time : 38.6838 secs
```



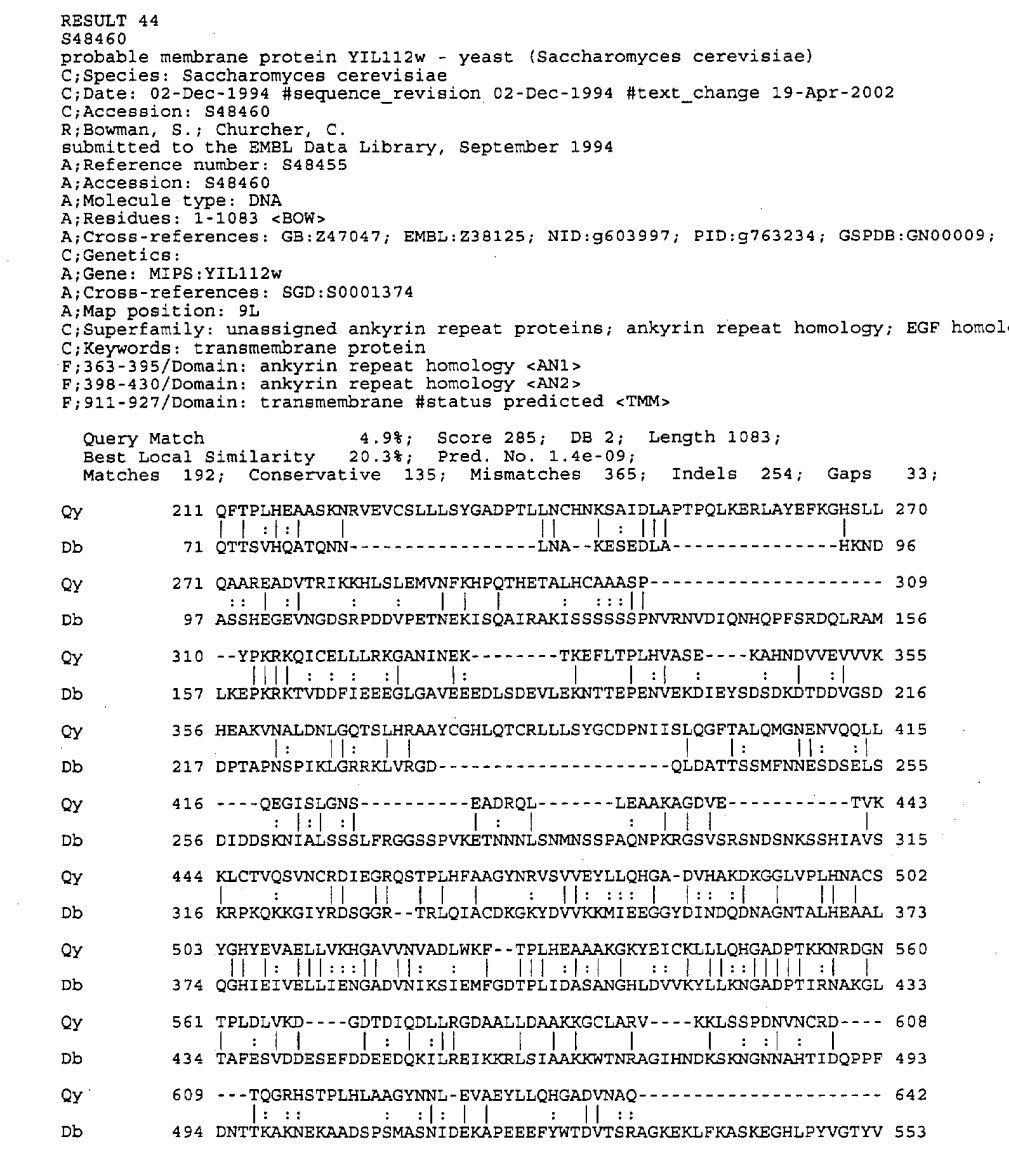
RESULT 43
T21720
hypothetical protein F34D10.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21720
R;Kershaw, J.
submitted to the EMBL Data Library, June 1994
A;Reference number: Z19464
A;Accession: T21720
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1295 <WIL>
A;Cross-references: EMBL:Z34799; PIDN:CAA84316.1; GSPDB:GN00021; CESP:F34D10.6
A;Experimental source: clone F34D10
C;Genetics:
A;Gene: CESP:F34D10.6
A;Map position: 3
A;Introns: 7/3; 46/3; 109/1; 204/2; 227/3; 266/3; 298/2; 336/1; 397/1; 429/3; 512/2; 597

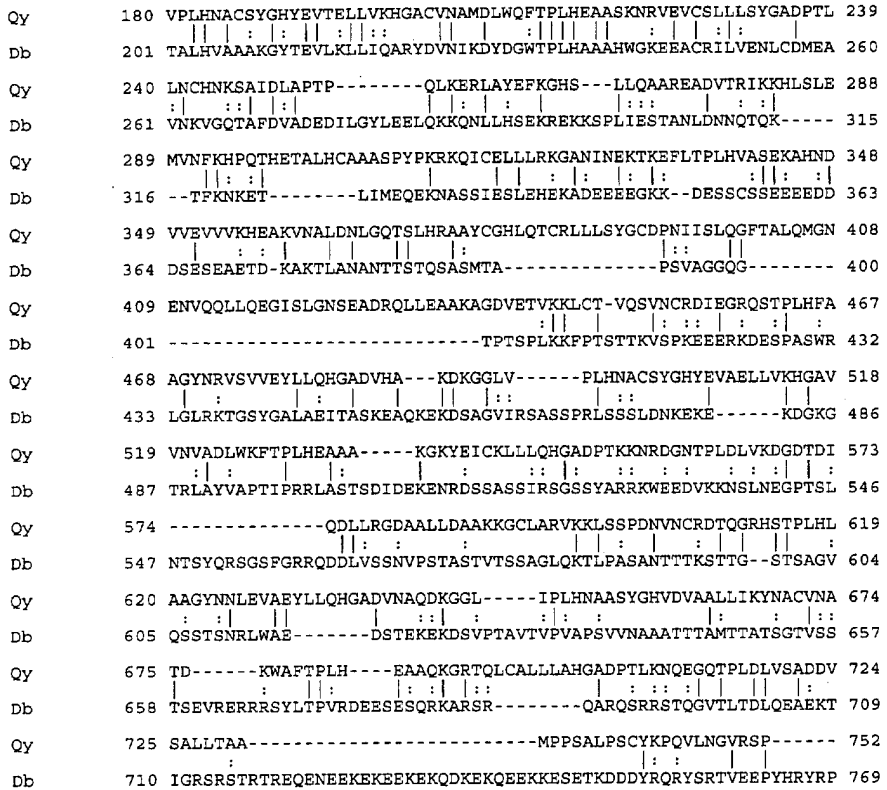
Query Match 5.0%; Score 287; DB 2; Length 1295;
Best Local Similarity 23.7%; Pred. No. 1.4e-09;
Matches 150; Conservative 93; Mismatches 228; Indels 162; Gaps 28;



RESULT 44
S48460
probable membrane protein YIL112w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
C;Accession: S48460
R;Bowman, S.; Churcher, C.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48455
A;Accession: S48460
A;Molecule type: DNA
A;Residues: 1-1083 <BOW>
A;Cross-references: GB:Z47047; EMBL:Z38125; NID:g603997; PID:g763234; GSPDB:GN00009; MIP
C;Genetics:
A;Gene: MIPS:YIL112w
A;Cross-references: SGD:S0001374
A;Map position: 9L
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: transmembrane protein
F;363-395/Domain: ankyrin repeat homology <AN1>
F;398-430/Domain: ankyrin repeat homology <AN2>
F;911-927/Domain: transmembrane #status predicted <TMM>

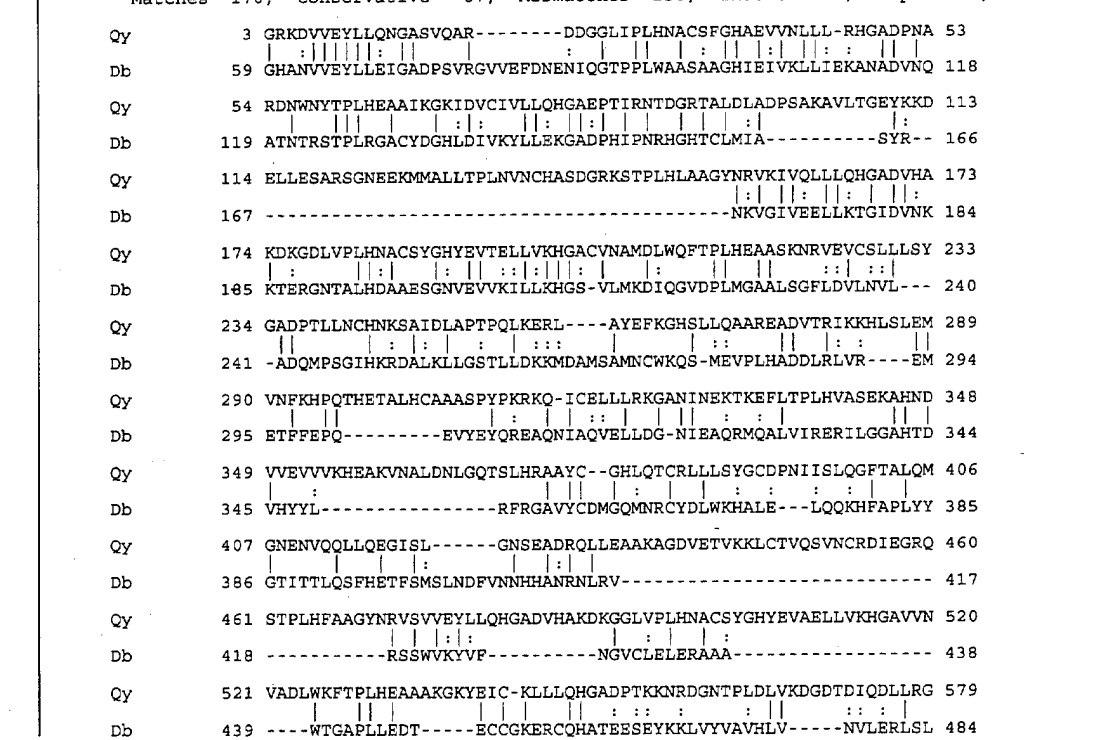
Query Match 4.9%; Score 285; DB 2; Length 1083;
Best Local Similarity 20.3%; Pred. No. 1.4e-09;
Matches 192; Conservative 135; Mismatches 365; Indels 254; Gaps 33;





RESULT 42
A34793
sex-determining protein fem-1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 29-Oct-1999
C;Accession: A34793; T30117
R;Spence, A.M.; Coulson, A.; Hodgkin, J.
Cell 60, 981-990, 1990
A;Title: The product of fem-1, a nematode sex-determining gene, contains a motif found ir
A;Reference number: A34793; MUID:90199879; PMID:2317869
A;Accession: A34793
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-656 <SPE>
A;Cross-references: GB:J03172; NID:g156309; PIDN:AAA28055.1; PID:g156310
R;Du, Z.; Le, T.T.
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of C. elegans cosmid F35D6.
A;Reference number: Z20739
A;Accession: T30117
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-656 <DUZ>
A;Cross-references: EMBL:U52001; PIDN:AAA96093.1; GSPDB:GN00022; CESP:fem-1
A;Experimental source: strain Bristol N2; clone F35D6
C;Genetics:
A;Gene: CESP:fem-1
A;Map position: 4
A;Introns: 25/3; 67/3; 88/1; 103/3; 139/1; 168/3; 290/3; 329/3; 378/2; 617/1
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F;47-79/Domain: ankyrin repeat homology <AN1>
F;88-119/Domain: ankyrin repeat homology <AN2>
F;122-154/Domain: ankyrin repeat homology <AN3>
F;155-187/Domain: ankyrin repeat homology <AN4>
F;188-219/Domain: ankyrin repeat homology <AN5>
F;220-252/Domain: ankyrin repeat homology <AN6>

Query Match 5.0%; Score 287.5; DB 2; Length 656;
Best Local Similarity 22.8%; Pred. No. 4.5e-10;
Matches 170; Conservative 87; Mismatches 256; Indels 233; Gaps 28;



```
Db        135 LLTHGAKINITDKVGCTPLHRAASVGKLEVCEFLIEEGAEIDATDKMGQTAL-------- 186

Qy        572 DIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEY 631

Db        187 ----------------------------MHSVICDDKQLKVSMD----------QVAFL 207

Qy        632 LLQHGADVNAQDKGGLIPLHNAAS 655

Db        208 LIRHGADVDVEDKEGYTVLGRATN 231


RESULT 40
A42024
transcription factor NF-kappa-B2, p100 splice form - human
N;Alternate names: lyt-10; major histocompatibility complex class I-specific transcripti
in NFKB2; transcription factor H2TF1
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Aug-2002
C;Accession: A42024; A41645; A48525
R;Bours, V.; Burd, P.R.; Brown, K.; Villalobos, J.; Park, S.; Ryseck, R.P.; Bravo, R.; K
Mol. Cell. Biol. 12, 685-695, 1992
A;Title: A novel mitogen-inducible gene product related to p50/p105-NF-kappa B participa
A;Reference number: A42024; MUID:92123193; PMID:1531086
A;Accession: A42024
A;Molecule type: mRNA
A;Residues: 1-900 <BOU>
A;Cross-references: GB:S76638; NID:g243420; PIDN:AAB21124.1; PID:g243421
A;Experimental source: peripheral blood T-cells
A;Note: sequence extracted from NCBI backbone (NCBIN:76638, NCBIP:76639)
R;Neri, A.; Chang, C.C.; Lombardi, L.; Salina, M.; Corradini, P.; Maiolo, A.T.; Chaganti
Cell 67, 1075-1087, 1991
A;Title: B cell lymphoma-associated chromosomal translocation involves candidate oncogen
A;Reference number: A41645; MUID:92103674; PMID:1760839
A;Accession: A41645
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-42,'A',44-143,'K',145-432,'EP',435-454,'HG',457-469,'A',471-663,'G',665-74
R;Potter, D.A.; Larson, C.J.; Eckes, P.; Schmid, R.M.; Nabel, G.J.; Verdine, G.L.; Sharp
J. Biol. Chem. 268, 18882-18890, 1993
A;Title: Purification of the major histocompatibility complex class I transcription fact
A;Reference number: A48525; MUID:93366804; PMID:8360178
A;Accession: A48525
A;Molecule type: protein
A;Residues: 'K',460-461,'X',463-469,'A';'KX',581-597;'KX',637-638,'X',640-645,'X',647 <P
A;Experimental source: HeLa cells
A;Note: sequence modified after extraction from NCBI backbone
C;Comment: For other splice forms, see PIR:S17233, PIR:I38609, PIR:A57034.
C;Comment: The full-length form of this protein is the MHC class I-specific transcriptic
.
C;Genetics:
A;Gene: GDB:NFKB2
A;Cross-references: GDB:128804; OMIM:164012
A;Map position: 10q24-10q24
C;Superfamily: human transcription factor NF-kappa-B2; ankyrin repeat homology; rel homo
C;Keywords: alternative splicing; DNA binding; nucleus; transcription factor; transcript
F;38-342/Domain: rel homology <REL>
F;310-313/Region: nuclear location signal
F;337-341/Region: nuclear location signal

  Query Match             5.1%;  Score 294.5;  DB 2;  Length 900;
  Best Local Similarity   26.9%;  Pred. No. 2.7e-10;
  Matches  133;  Conservative   58;  Mismatches  182;  Indels  121;  Gaps   19;

Qy        406 MGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLH 465

Db        453 LAQRSARALLDYGVT-----ADRRALLAGQ-------RHLLTAQDEN-------GDTPLH 493

Qy        466 FAAGYNRVSVVE---YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVA 522

Db        494 LAIIHGQTSVIEQIVYVIHHAQDL----------------------------GVVNLT 523

Qy        523 DLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAA 582

Db        524 NHLHQTPLHLAVITGQTSVVSFLLRVGADPALLDRHGDSAMHLALRAGAGAPELLR---A 580

Qy        583 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642

Db        581 LLQSGAP---------AVPQLLHMPDFEGLY--PVHLAVRARSPECLDLLVDSGAEVEAT 629

Qy        643 DK-GGLIPLHNAASYGHVD-VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 700

Db        630 ERQGGRTALHLATEMEELGLVTHLVTKLRANVNARTFAGNTPLHLAAGLGYPTLTRLLLK 689

Qy        701 HGADPTLKNQE----------------------------GQTPLDLVSADDVSALLTAA 731

Db        690 AGADIHAENEEPLCPLPSPPTSDSDSDSEGPEKDTRSSFRGHTPLDLTCSTKVKTLLLNA 749

Qy        732 M-----PPSALPSCYKPQVLNGVRSPGATA----DALSSGPSSPSSLSAASSLDNLSGSF 782

Db        750 AQNTMEPPLTPPSPAGPGL-----SLGDTALQNLEQLLDGPEAQGSWAELAERLGLRSLV 804

Qy        783 SELSSVVSSSGTEGASSLEKKEVPGVDFS-ITQFVRNLGLEHLMDIFEREQITLDVLVEM 841

Db        805 DTYRQTTSPSG----SLLRSYELAGGDLAGLLEALSDMGLEEGVRLL-RGPETRDKLPST 859

Qy        842 GHKELKEIGINAYG 855

Db        860 --TEVKE--DSAYG 869


RESULT 41
A55142
myosin-light-chain-phosphatase (EC 3.1.3.53) large chain, long form - chicken
N;Alternate names: smooth muscle protein phosphatase 1M chain M110
C;Species: Gallus gallus (chicken)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: A55142; S51023
R;Shimizu, H.; Ito, M.; Miyahara, M.; Ichikawa, K.; Okubo, S.; Konishi, T.; Naka, M.; Ta
J. Biol. Chem. 269, 30407-30411, 1994
A;Title: Characterization of the myosin-binding subunit of smooth muscle myosin phosphat
A;Reference number: A55142; MUID:95074044; PMID:7982954
A;Accession: A55142
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1004 <SHI>
A;Cross-references: GB:D37985; NID:g633037; PIDN:BAA07201.1; PID:d1007779; PID:g633038;
R;Chen, Y.H.; Chen, M.X.; Alessi, D.R.; Campbell, D.G.; Shanahan, C.; Cohen, P.; Cohen,
FEBS Lett. 356, 51-55, 1994
A;Title: Molecular cloning of cDNA encoding the 110 kDa and 21 kDa regulatory subunits c
A;Reference number: S51022; MUID:95080427; PMID:7988720
A;Accession: S51023
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 681-1004 <CHE>
A;Cross-references: GB:S74623; NID:g802102; PIDN:AAB32730.1; PID:g802103
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C;Keywords: alternative splicing; phosphoric monoester hydrolase; smooth muscle; tandem
F;72-104/Domain: ankyrin repeat homology <AN1>
F;105-137/Domain: ankyrin repeat homology <AN2>

  Query Match             5.0%;  Score 290;  DB 2;  Length 1004;
  Best Local Similarity   19.9%;  Pred. No. 6.1e-10;
  Matches  220;  Conservative  155;  Mismatches  423;  Indels  306;  Gaps   40;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         51 GDTEEVLRLLERGADINYANVDGLTALHQACIDDNVDMVKFLVENGANINQPDNEGWIPL 110

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSA-KAVLTGEYKKDEL-LESAR 120

Db        111 HAAASCGYLDIAEYLISQGAHVGAVNSEGDTPLDIAEEEAMEELLQNEVNRQGVDIEAAR 170

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADV-HAKDKGDL 179

Db        171 K-EEERIM-----------LRDARQ----WLNSG-----------HINDVRHAKSGG-- 200
```

```
A;Residues: 1-1016 <WI2>
A;Cross-references: EMBL:Z47809; PIDN:CAA87782.1; GSPDB:GN00020; CESP:C06C3.1
A;Experimental source: clone F42A8
C;Genetics:
A;Gene: CESP:C06C3.1
A;Map position: 2
A;Introns: 27/3; 94/3; 279/3; 352/2; 462/2; 523/3; 569/2; 657/2; 718/3; 766/3; 833/3; 89
C;Keywords: phosphoric monoester hydrolase

  Query Match             5.2%;  Score 300.5;  DB 2;  Length 1016;
  Best Local Similarity   28.5%;  Pred. No. 1.4e-10;
  Matches   94;  Conservative   48;  Mismatches  121;  Indels   67;  Gaps   11;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              | :: || ||  ||::      ||  || :      |:|  |   ||| ||:||  :|||
Db         66 GDEEEVEELLNKGANINTCTVDGLTALHQSVIDSKPEMVRFLCEKGADVNAQDNEGWTPL 125

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLA-DPSAKAVLTGEYKKDEL-LESAR 120
              | ||  | : :  | ||||: :| |:|   |||||  |   :  |  :||:  : |:: |
Db        126 HAAACCGNVAIVRYLCQHGADLSIVNSDKELALDLAVDEQCRDYLDDDYKRQMIDLDACR 185

Qy        121 SGNEEKMMALLTPLNVNCHASDGR-KSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDL 179
                ::::  :|   :||   | |  :  | |   |
Db        186 ---DQELQTMLK--DVNMWMSQGEYRDIPHHRTGG------------------------ 215

Qy        180 VPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGAD--- 236
                :| |   |: :: |||:| |  | | |:  :|||| ||     : | :||  ||:
Db        216 TAMHVAAGRGYTQLLELLIKAGGNVRAQDVEGWTPLHAAAHWAERDACKILLENGAELSD 275

Qy        237 -----PTLLNCHNKSAIDL----------------------APTPQLKERLAYEFKGHS- 268
                  :|   :|  ||                        ||  |:|:  :    |
Db        276 LTFTGADVLGVADKECIDYLVELADTVKVQNKRKSPGSGSQPPTSILQEK-NHRMPSHEE 334

Qy        269 -LLQAAREADVTRIKKHLSLEMVNFKHPQT 297
               :| : |: |:    :|   | ::  || |
Db        335 HVLTSERKRDLQHKDQHSENEPLH-SHPST 363


RESULT 38
T27499
hypothetical protein ZC15.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27499
R;Basham, V.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z20377
A;Accession: T27499
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-633 <WIL>
A;Cross-references: EMBL:Z93396; NID:e1188462; PIDN:CAB07714.1; GSPDB:GN00023; CESP:ZC15
A;Experimental source: clone ZC15
C;Genetics:
A;Gene: CESP:ZC15.7
A;Map position: 5
A;Introns: 15/3; 69/2; 100/3; 148/3; 194/1; 305/1; 404/3; 520/3

  Query Match             5.2%;  Score 299.5;  DB 2;  Length 633;
  Best Local Similarity   23.8%;  Pred. No. 7.5e-11;
  Matches  161;  Conservative   90;  Mismatches  257;  Indels  169;  Gaps   26;

Qy         62 LHEAAIK-GKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              | : ||| |  ||      |  |  | :  ||  |         | :|    |  : :   |
Db         25 LEKQAIKNGSTDV-----DHFWE--ILDDKGRNVL------FHAAITDNLKNFQFIMKTR 71

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQ------------H 167
                :| :: |          :|  |:| || |  ||  ::|: ::|              |
Db         72 RVKKEALLRL----------TDYNKATVLHWATQYNCKQVVKEIVQTFDKLNPEDLKILH 121

Qy        168 GADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQF-------TPLHEAAS 220
              |  : |||   : ||| | :     :: :: |:       |:|:           :||| ||
Db        122 GL-ILAKDSESVTPLHIAATKQDTKILKIFVEILKTPKVMELFSIVKDKRDRSPLHYAAC 180

Qy        221 KNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYE----FKGHSLLQAAREA 276
              | :|   :||    ||                |   |   :::           | :| ||
Db        181 KVNLEALRILLF--VDP-------NGGPDFGFTVDQRDKFGIIPLMCAVGVNLPQAIPVI 231

Qy        277 DVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLT 336
                  || :|    |     :    |||| | |:   :  :  :||:  |::::     |  |
Db        232 RFLEKKKPVSKTRQN----RDGMTALHIAVAA---RNLEAVQLLIELGSSVDLVDNEQRT 284

Qy        337 PLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNII 396
              ||| |:|: : ::|: :: : |: :  |::| |  | ||
Db        285 PLHYAAEQGYPEIVKFLLCNGARNSTRDHIGATPAHYAA---------------------- 323

Qy        397 SLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDI 456
                    |   | :: |  |                                : |:|  |
Db        324 --------QFSVECLKILFAES--------------------------KITEVN--DN 345

Qy        457 EGRQSTPLHFAAGYNRVSVVEYLLQH--GADVHAKDKGGLVPLHNACSYGHYEVAELLVK 514
              |||    | :|     : |: ||:|          | || |   || |   || :| ::|
Db        346 EGRSC--LMWAVCAGNIEVINYLIQREDAPKRAACDKNGYTALHLAAMVGHEKVCKILTN 403

Qy        515 ---------HGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL 565
                      |  ::  |    | || |: :|  :: : |:  ||:    :  | | :
Db        404 QPTSIFPYFQGWSLSERDNHSNTALHLASGRGHTDVLRCLVASGANMNDVDEVGRTAV-- 461

Qy        566 VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLS----SPDNVNCR---DTQGRHSTPLH 618
                      : | |   |      |: :|:|      :  :|| |   |  | || |||
Db        462 -------FWACMGGQAHTLH-----CM--IKELGFEWRTSGSVNTRPKCDKNGR--TALH 505

Qy        619 LAAGYNNLEVAEYLLQHGAD-------VNAQDKGGLIPLHNAASYGHVDVAALLIKYNAC 671
              ||  ::  :   |      |      | || |   || |     :|  |:      :
Db        506 AAAVESSACINVLLSIEKEDNFLSSPLVGWLDKNGETALHEACLASRIDCILSLLNGGSA 565

Qy        672 VNATDKWAFTPLHEAAQ 688
              |||  :   |||  | |
Db        566 VNAFGRPQRTPLDCAMQ 582


RESULT 39
D84448
probable ankyrin [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 08-Sep-2002
C;Accession: D84448
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84448
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-247 <STO>
A;Cross-references: GB:AE002093; NID:g4335756; PIDN:AAD17433.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g03430
A;Map position: 2
C;Superfamily: hypothetical protein YGR232w; ankyrin repeat homology

  Query Match             5.2%;  Score 299;  DB 2;  Length 247;
  Best Local Similarity   30.7%;  Pred. No. 1.8e-11;
  Matches   81;  Conservative   38;  Mismatches   87;  Indels   58;  Gaps    5;

Qy        401 FTALQMGNENV-----QQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQS----V 451
              | | :  |: ::     ::  | : ::   | :      |  ||   |  : || | :      :
Db         17 FKAAEWGDSSLFMSLSEEQLSKSLNFRNEDGRSLLHVAASFGHSQIVKLLSSSDEAKTVI 76

Qy        452 NCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL 511
              | :| ||    ||| ||      :|| ||  ||||:||: ||    || | | |  |:|:|
Db         77 NSKDDEG--WAPLHSAASIGNAELVEVLLTRGADVNAKNNGGRTALHYAASKGRLEIAQL 134

Qy        512 LVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT 571
```

```
C;Accession: A59288
R;Patel, K.G.; Liu, C.; Cameron, P.L.; Cameron, R.S.
submitted to GenBank, November 1999
A;Description: Identification of a Novel Mammalian Myosin Class, XVI, in Developing Brai
A;Reference number: A59288
A;Accession: A59288
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1322 <PAT>
A;Cross-references: GB:AF209114; PIDN:AAF20150.1
A;Experimental source: strain Sprague-Dawley; clone KP4; cell type type 1 astrocyte
C;Superfamily: myosin motor domain homology
F;404-1132/Domain: myosin motor domain homology <MMO>

  Query Match             5.3%;  Score 306;  DB 2;  Length 1322;
  Best Local Similarity   23.9%;  Pred. No. 9.3e-11;
  Matches  151;  Conservative   98;  Mismatches  226;  Indels  158;  Gaps   24;

Qy        411 VQQLLQEGISLGNSEADRQLLEAAKAGDVETV---KKLCTVQSVNCRDIEGRQSTPLHF- 466
              :  | | | : ||  :   :|:  :  ::    :|:   |    : |:  :|  : |
Db          3 IDQCLLESLPLGQRQ---RLVRRMRCEQIKAYYEREKVFQKQEGLLKRIKPGKSQKVRFG 59

Qy        467 -------AAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVV 519
                    |  ::    |  ||: ||| |     |    ||    | :  :||:|:  |  |
Db         60 LADMIQDAIIHHHDKEVLQLLKEGADPHTLVSSGGSLLHLCARYDNVFIAEVLIDRGVNV 119

Qy        520 NVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRG 579
              |  |   : |:| | |    :|  ||:  ||:  :: ::|| |||   :|      ||
Db        120 NHQDEDFWAPMHIACACDNPDIVLLLILAGANVLLQDVNGNIPLDYAVEGTESSAILL-- 177

Qy        580 DAALLD--AAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGA 637
                 | ||         | ::|                  || :      |    : |  |
Db        178 --AYLDENGVDLNSLRQIKL---------------QRPLSM------LTDVRHFLSSGG 213

Qy        638 DVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCAL 697
              ||| ::  |:  || | : |: :|  ||:::   :|  |   :|||| ||: |:| |  |
Db        214 DVNEKNDDGVTLLHMACASGYKEVVLLLLEHGGDLNGMDDGYWTPLHLAAKYGQTTLVKL 273

Qy        698 LLAHGADPTLKNQEGQTPLDLVSADDVSALLTAA------------MPPSALPSCYK-- 742
              |||| |:| | |  |: | |: ::: :  :|  |              :|  |     |:
Db        274 LLAHQANPHLVNCNGEKPSDIAASESIEEMLLKAEIAWEERMKESPSVPSLAQEELYEEI 333

Qy        743 ----PQV---LNGVRSPGATADAL------------------SSGPSSPSSLSAASSLD 776
                  |::   |: :  | |  |:|                  ||   | :|::|  :
Db        334 LHDLPELSSKLSPLVLPIAKQDSLLEKDIMFKDTTKGLCNQESQDGPPETSMVSSSSKPE 393

Qy        777 NL-------SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFE 829
               :       |   : || :  ||         | ::  | |  |  |: :: |  |:: |:
Db        394 QVQLTPPAPSDDLATLSELNDSS-----LLYEIQKRFGND-QIHTFIGDIFL--LVNPFK 445

Qy        830 REQITLDVLVEMGHKELKEIGINAYGHR----------------HKLIKGVER-----LI 868
                  |   |: :|          ::  | |                  |:| :  ||      ::
Db        446 ELPIYSTVVSQM--------YLSPTGQRSPSLPPHLFSCAERAFHRLFQ--ERRPQNIIL 495

Qy        869 SGQQG---------LNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHA 919
              ||::|         :  :||   | | |:        |  |:      :::          |||
Db        496 SGERGSGKTQASKQIMKHLTSRASSSCTMF------DSRFKHAICILEAF-------GHA 542

Qy        920 GGIFNRYNILKIQKVCNKKLWE-RYTHRRKEVS 951
                  |  :   ||        || ::  ||| |:
Db        543 KTTLNNVSSCLIQ------YWELQFCQRRKHVT 569


RESULT 36
B45771
2-5A-dependent RNAase - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
C;Accession: B45771
R;Zhou, A.; Hassel, B.A.; Silverman, R.H.
Cell 72, 753-765, 1993
A;Title: Expression cloning of 2-5A-dependent RNAase: a uniquely regulated mediator of i
A;Reference number: A45771; MUID:93201598; PMID:7680958
A;Accession: B45771
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-679 <ZHO>
A;Cross-references: GB:L10382; NID:g293264; PIDN:AAA37117.1; PID:g293265
A;Note: sequence extracted from NCBI backbone (NCBIP:127769)
C;Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homolog
F;58-90/Domain: ankyrin repeat homology <AN22>
F;124-156/Domain: ankyrin repeat homology <AN21>

  Query Match             5.3%;  Score 305;  DB 2;  Length 679;
  Best Local Similarity   27.5%;  Pred. No. 3.8e-11;
  Matches  148;  Conservative   69;  Mismatches  232;  Indels   90;  Gaps   23;

Qy        425 EADRQLLEAAKAGDVETVKKLCT--VQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQH 482
              | |  |::| : |||  |::|      :  | |  |  ||||  |    || :|  || |
Db         24 EDDSSLIKAVQKGDVVRVQQLLEKGADANACEDTWG--WTPLHNAVQAGRVDIVNLLLSH 81

Qy        483 GADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEIC 542
              ||| | : | |  |    |   | :: |:|:  || ||  |   ||   ||| :|  |
Db         82 GADPHRRKKNGATPFIIAGIQGDVKLLEILLSCGADVNECDENGFTAFMEAAERGNAEAL 141

Qy        543 KLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDA-ALLDAAKKGCL--ARVKKLS 599
              : |   ||:          ::| :    | : | :| | ||: ||:|| |    |:
Db        142 RFLFAKGAN-----------VNLRRQTTKDKRRLKQGGATALMSAAEKGHLEVLRILLND 190

Qy        600 SPDNVNCRDTQGRHS---TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASY 656
                  |: ||  ||::   | |:    |  |:    |:|||||| : : |  ||  |
Db        191 MKAEVDARDNMGRNALIRTLLNWDC-ENVEEITSILIQHGADVNVRGERGKTPLIAAVER 249

Qy        657 GHVDVAALLIKYNAC-VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTP 715
              |  :  :|:       ::| |    | |  |  |   ::  |||  |||
Db        250 KHTGLVQMLLSREGINIDARDNEGKTALLIAVDKQLKEIVQLLLEKGAD----------- 298

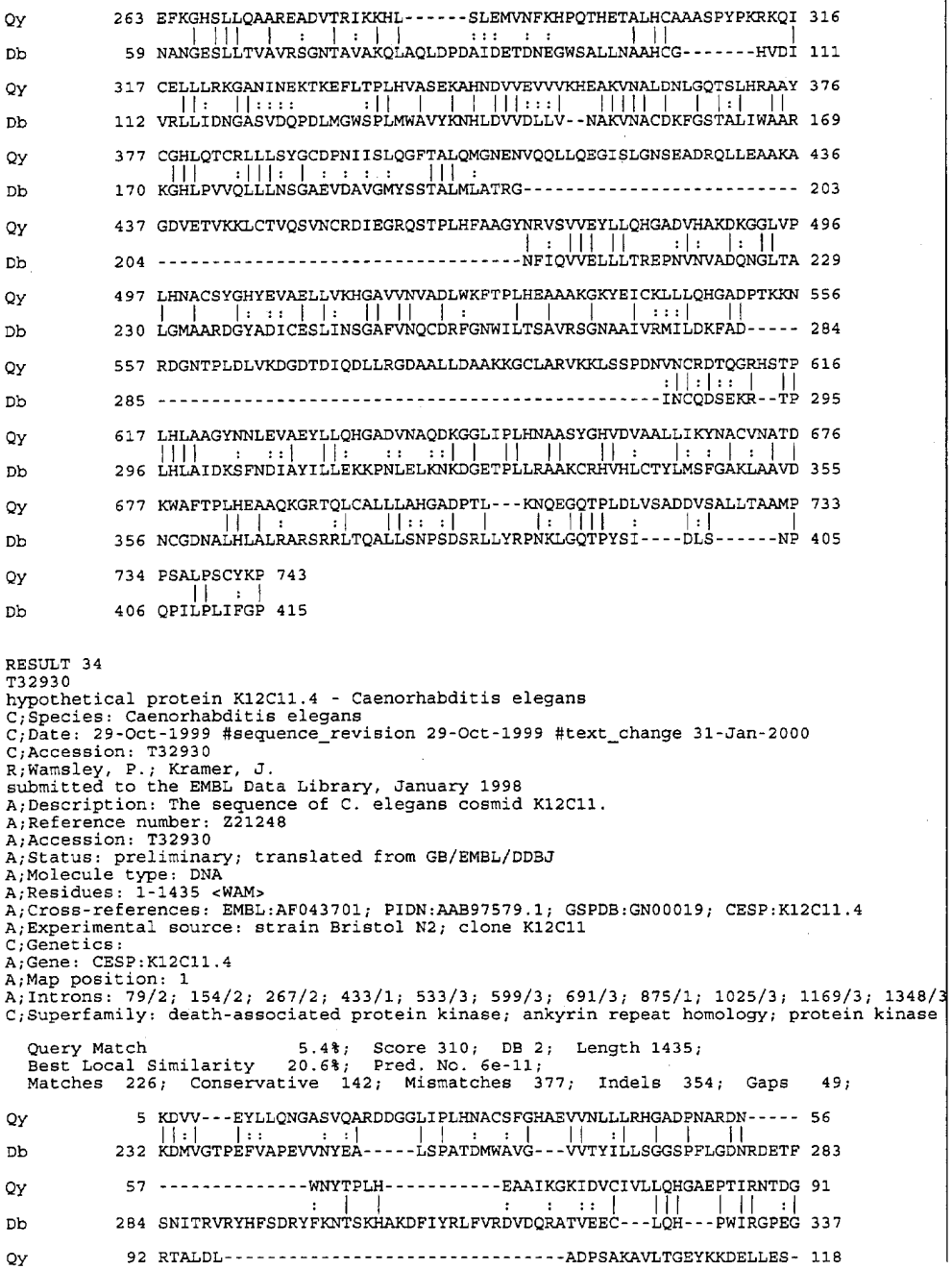
Qy        716 LDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSL 775
                  ||:  :          |      | | |     | |        |  ||| |    ::|
Db        299 ----KCDDL-VWIARRNHDYHLVKLLLPYVANPDTDPPA-------GDWSPHSSRWGTAL 346

Qy        776 DNLSG----SFSELSSVVSS----SGT-EGASSL---EKKEVPGVDF--------SITQF 815
                :|         :|   :       :|| |||  |   : :||    |
Db        347 KSLHSMTRPMIGKLKIFIHDDYKIAGTSEGAVYLGIYDNREVAVKVFRENSPRGCKEVSC 406

Qy        816 VRNLGLEH--LMDIFERE------------QITLDVLVEMGHKELKEIGINAYGHR--H 858
              :|: | :|  |:  : ||             : ||:  : :  :|  | | : :  |
Db        407 LRDCG-DHSNLVAFYGREDDKGCLYVCVSLCEWTLEEFLRLPREEPVENGEDKFAHSILL 465

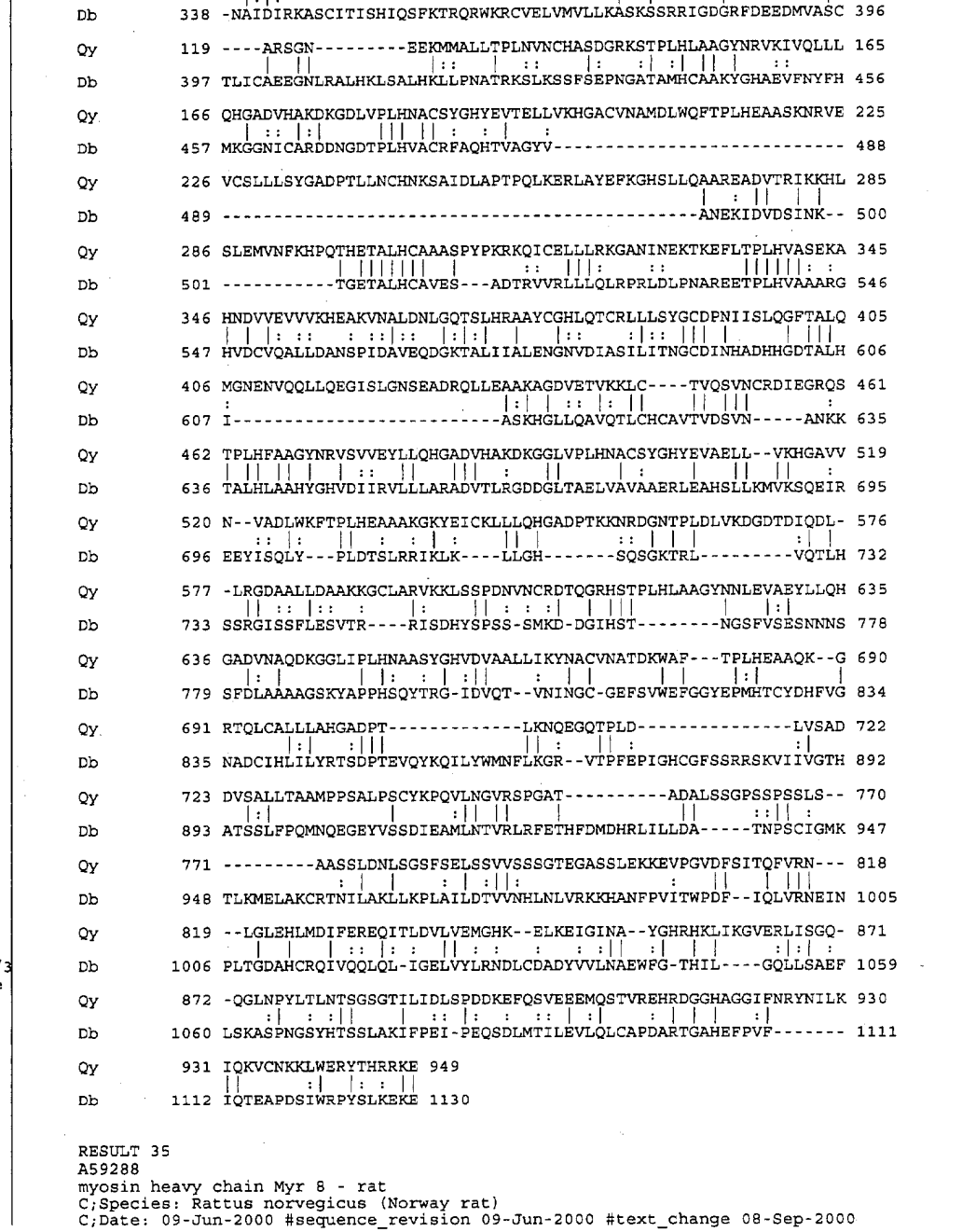
Qy        859 KLIKGVERL-ISG--QQGLNPYLTLNTSGSGTILIDLSPDDK---EFQSVEEEMQSTVR 911
              : :||::| : |  | | |  |  |    :  | | |  :::   |
Db        466 SIFEGVQKLHLHGYSHQDLQPQNILIDSKKAVRLADFDQSIRWMGESQMVRRDLEDLGR 524


RESULT 37
T19006
ankyrin related protein C06C3.1 - Caenorhabditis elegans
N;Contains: myosin-light-chain-phosphatase (EC 3.1.3.53)
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
C;Accession: T19006; T22086
R;Berks, M.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19058
A;Accession: T19006
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1016 <WIL>
A;Cross-references: EMBL:Z36719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:C06C3.1
A;Experimental source: clone C06C3
R;Matthews, P.
submitted to the EMBL Data Library, January 1995
A;Reference number: Z19510
A;Accession: T22086
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
```



RESULT 34
T32930
hypothetical protein K12C11.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C;Accession: T32930
R;Wamsley, P.; Kramer, J.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid K12C11.
A;Reference number: Z21248
A;Accession: T32930
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1435 <WAM>
A;Cross-references: EMBL:AF043701; PIDN:AAB97579.1; GSPDB:GN00019; CESP:K12C11.4
A;Experimental source: strain Bristol N2; clone K12C11
C;Genetics:
A;Gene: CESP:K12C11.4
A;Map position: 1
A;Introns: 79/2; 154/2; 267/2; 433/1; 533/3; 599/3; 691/3; 875/1; 1025/3; 1169/3; 1348/3
C;Superfamily: death-associated protein kinase; ankyrin repeat homology; protein kinase

Query Match 5.4%; Score 310; DB 2; Length 1435;
Best Local Similarity 20.6%; Pred. No. 6e-11;
Matches 226; Conservative 142; Mismatches 377; Indels 354; Gaps 49;

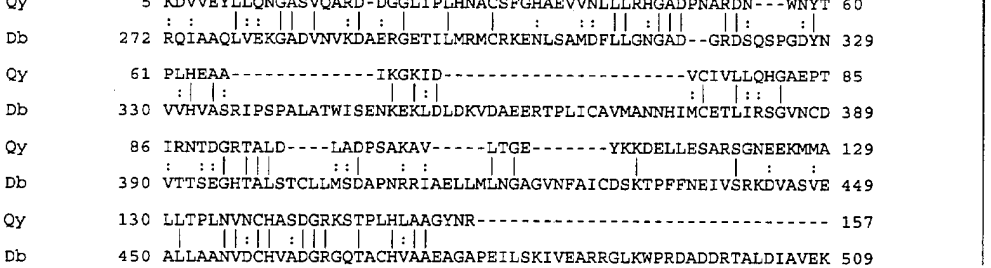


RESULT 35
A59288
myosin heavy chain Myr 8 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000



RESULT 32
T21280
hypothetical protein F22G12.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21280
R;Lennard, N.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19400
A;Accession: T21280
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1107 <WIL>
A;Cross-references: EMBL:Z92831; PIDN:CAB07368.1; GSPDB:GN00019; CESP:F22G12.4
A;Experimental source: clone F22G12
C;Genetics:
A;Gene: CESP:F22G12.4
A;Map position: 1
A;Introns: 23/3; 153/2; 216/2; 268/1; 304/1; 363/2; 521/1; 562/2; 643/3; 738/3; 770/3; 8

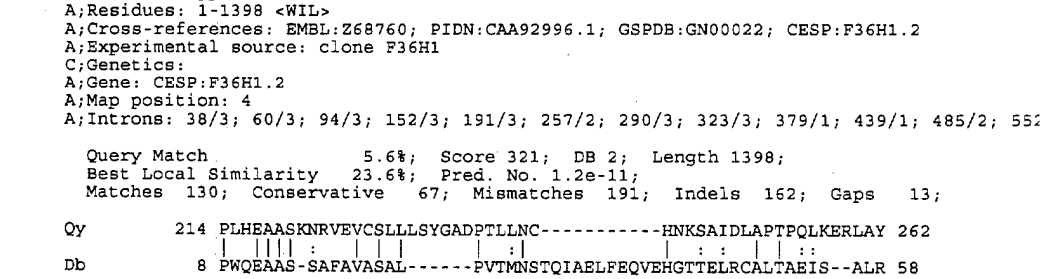
Query Match 5.8%; Score 336.5; DB 2; Length 1107;
Best Local Similarity 21.7%; Pred. No. 8.6e-13;
Matches 185; Conservative 119; Mismatches 311; Indels 239; Gaps 32;





RESULT 33
T21884
hypothetical protein F36H1.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21884
R;Kershaw, J.
submitted to the EMBL Data Library, January 1996
A;Reference number: Z19482
A;Accession: T21884
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1398 <WIL>
A;Cross-references: EMBL:Z68760; PIDN:CAA92996.1; GSPDB:GN00022; CESP:F36H1.2
A;Experimental source: clone F36H1
C;Genetics:
A;Gene: CESP:F36H1.2
A;Map position: 4
A;Introns: 38/3; 60/3; 94/3; 152/3; 191/3; 257/2; 290/3; 323/3; 379/1; 439/1; 485/2; 552

Query Match 5.6%; Score 321; DB 2; Length 1398;
Best Local Similarity 23.6%; Pred. No. 1.2e-11;
Matches 130; Conservative 67; Mismatches 191; Indels 162; Gaps 13;



```
Qy        505 HYEVAELLVKHGAVVNVADLWKFTPLHEAA-------AKGKYEICKLLLQHGAD--PTKK 555
Db        375 CLRTMQLLLDHGANIDA-----YIATHPTAFPATIMFAMKCLSLLKFLMDLGCDGEPCFS 429

Qy        556 NRDGNTP 562
Db        430 CLYGNGP 436
```

RESULT 30
AC2508
hypothetical protein alr7243 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AC2508
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2508
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-606 <KUR>
A;Cross-references: GB:BA000020; PIDN:BAB78327.1; PID:g17135781; GSPDB:GN00180
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7243
A;Genome: plasmid

Query Match 6.0%; Score 348.5; DB 2; Length 606;
Best Local Similarity 25.2%; Pred. No. 5.8e-14;
Matches 126; Conservative 68; Mismatches 147; Indels 159; Gaps 17;

```
Qy        305 AAASPYPKR-----KQICELLLRK-------------------GANINEKTKEFLTPLH 339
Db         33 AVESKIPKKTSNSISPFCKVLSEKFSPFGAAPDLSVALGKAIAEGGNLNQ-------PCD 85

Qy        340 VASEK---------AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ-TCRLLLSY 389
Db         86 IFGEKWLPLNFLLSTDEELSQRLIQTGVNVNAKDGKGDTPLH---YLGKSEKNARLLLSR 142

Qy        390 GCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQ 449
Db        143 GAKVNV-----------GNQN-----------GN------------------------- 154

Qy        450 SVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVA 509
Db        155 ------------TPLHNLFGQKTAYVIKLLIDRGAKVNARNQEQMTPLHFAANSGKADIT 202

Qy        510 ELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDG 569
Db        203 ELLISRGGDVNARSTQNWTPLHYGAS--NLEVAKKLIQAGADLTIQNREG---------- 250

Qy        570 DTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVA 629
Db        251 ----------AVIHSSSLEPAVLKL-LLDRGVNVNLRNRQGQ--TPLHIHRFQPAL--V 294

Qy        630 EYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQK 689
Db        295 KLLLARKAQVNLRDAQGKTPLYDV----NLEVARQLVAANADLNVQDNLGRTPLHQAVLE 350

Qy        690 GR----TQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQV 745
Db        351 EQSFFGSELVTLFLSKNARVDLKDKQGKTALDLAR-----------------QLNKTQI 392

Qy        746 LNGVRSPGATADALSSGPSS 765
Db        393 INLLERHMATKNSPGSGTST 412
```

RESULT 31
I37275
death-associated protein kinase (EC 2.7.1.-) - human
N;Alternate names: calmodulin-dependent protein kinase homolog; DAP kinase
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 10-Jul-1998
C;Accession: I37275; S39269
R;Deiss, L.P.; Feinstein, E.; Berissi, H.; Cohen, O.; Kimchi, A.
Genes Dev. 9, 15-30, 1995
A;Title: Identification of a novel serine/threonine kinase and a novel 15-kD protein as
A;Reference number: A55614; MUID:95129831; PMID:7828849
A;Accession: I37275
A;Molecule type: mRNA
A;Residues: 1-1423 <RES>
A;Cross-references: EMBL:X76104; NID:g434846; PID:g434847
C;Genetics:
A;Gene: GDB:DAPK1; DAPK
A;Cross-references: GDB:555932; OMIM:600831
A;Map position: 9q34.1-9q34.1
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
A;Pathway: apoptosis
A;Note: activity is calmodulin dependent
C;Superfamily: death-associated protein kinase; ankyrin repeat homology; protein kinase
C;Keywords: apoptosis; ATP; calmodulin binding; phosphotransferase; serine/threonine-spe
F;11-267/Domain: protein kinase homology <KIN>
F;19-27/Region: protein kinase ATP-binding motif
F;285-308/Region: calmodulin binding #status predicted
F;370-402/Domain: ankyrin repeat homology <AN1>
F;403-435/Domain: ankyrin repeat homology <AN2>
F;436-468/Domain: ankyrin repeat homology <AN3>
F;470-502/Domain: ankyrin repeat homology <AN4>
F;503-535/Domain: ankyrin repeat homology <AN5>
F;536-568/Domain: ankyrin repeat homology <AN6>
F;569-601/Domain: ankyrin repeat homology <AN7>
F;602-634/Domain: ankyrin repeat homology <AN8>
F;42,64,139,141/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 6.0%; Score 345; DB 1; Length 1423;
Best Local Similarity 22.0%; Pred. No. 3.7e-13;
Matches 207; Conservative 119; Mismatches 355; Indels 258; Gaps 33;

```
Qy         58 NYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLE 117
Db        182 NYEPLGLEADMWSIGVITYILLSGASPFLGDTKQET---LANVSA---VNYEF-EDEYFS 234

Qy        118 SARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKG 177
Db        235 NTSALAKDFIRRLL--------VKDPKKR-----------MTIQDSLQH-PWIKPKD-- 271

Qy        178 DLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADP 237
Db        272 --------------TQQALSRKASAVNMEKFKKF-----AARKKWKQSVRLI------- 304

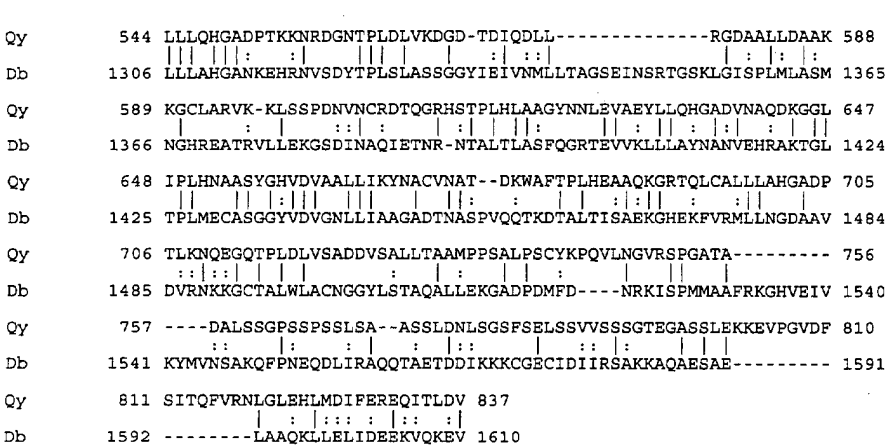
Qy        238 TLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQT 297
Db        305 ----------------SLCQRLSRSFLSRSNMSVARSDDTLDEEDSFVMKAI------- 340

Qy        298 HETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHE 357
Db        341 ----IHAINDDNVPGLQHLLGSL--SNYDVNQPNKHGTPPLLIAAGCGNIQILQLLIKRG 394

Qy        358 AKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQE 417
Db        395 SRIDVQDKGGSNAVYWAARHGHVDTLKFLSENKCPLDVKDKSGEMALHV----------- 443

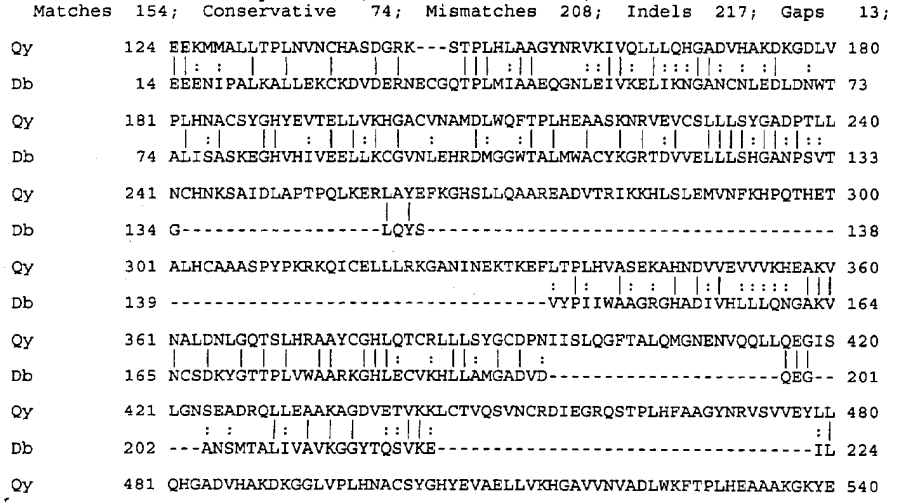
Qy        418 GISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVE 477
Db        444 --------------AARYGHADVAQVTCAASAQIPISRTKEEETPLHCAAWHGYYSVAK 488

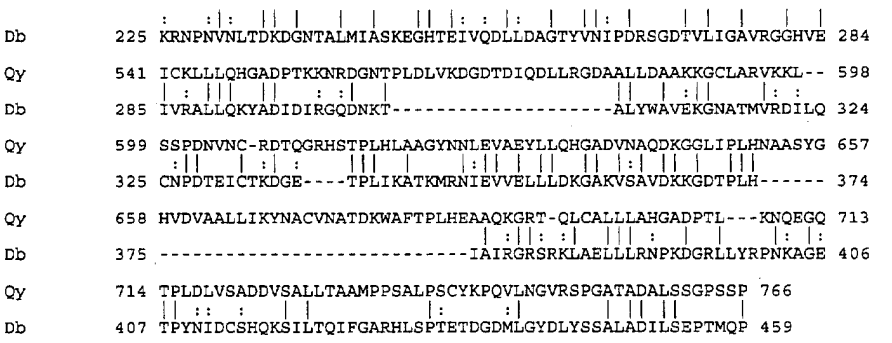
Qy        478 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 537
Db        489 ALCEAGCNVNIKNREGETPLLTASARGYHDIVECLAEHGADLNACDKDGHIALHLAVRRC 548
```



RESULT 28
T43458
hypothetical protein DKFZp434F0621.1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T43458
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z22517
A;Accession: T43458
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1031 <AAA>
A;Cross-references: EMBL:AL133620
A;Experimental source: adult testis; clone DKFZp434F0621
C;Genetics:
A;Note: DKFZp434F0621.1

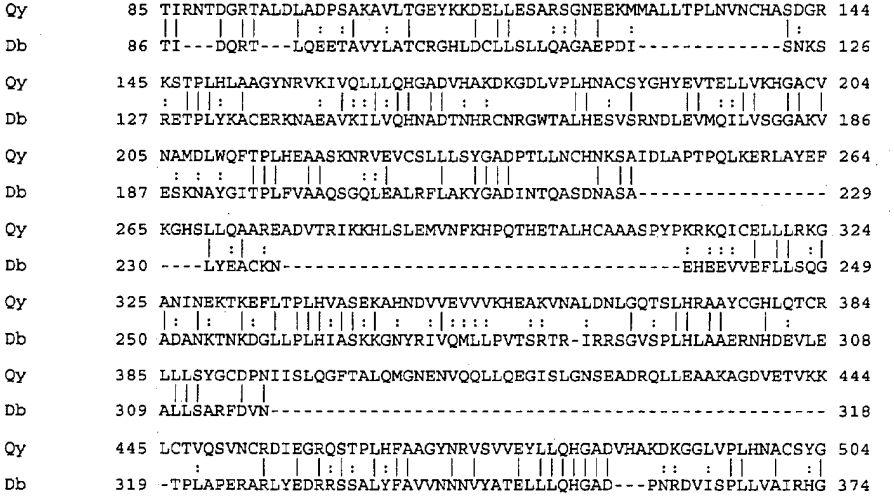
Query Match 6.3%; Score 363.5; DB 2; Length 1031;
Best Local Similarity 23.6%; Pred. No. 1.5e-14;
Matches 154; Conservative 74; Mismatches 208; Indels 217; Gaps 13;

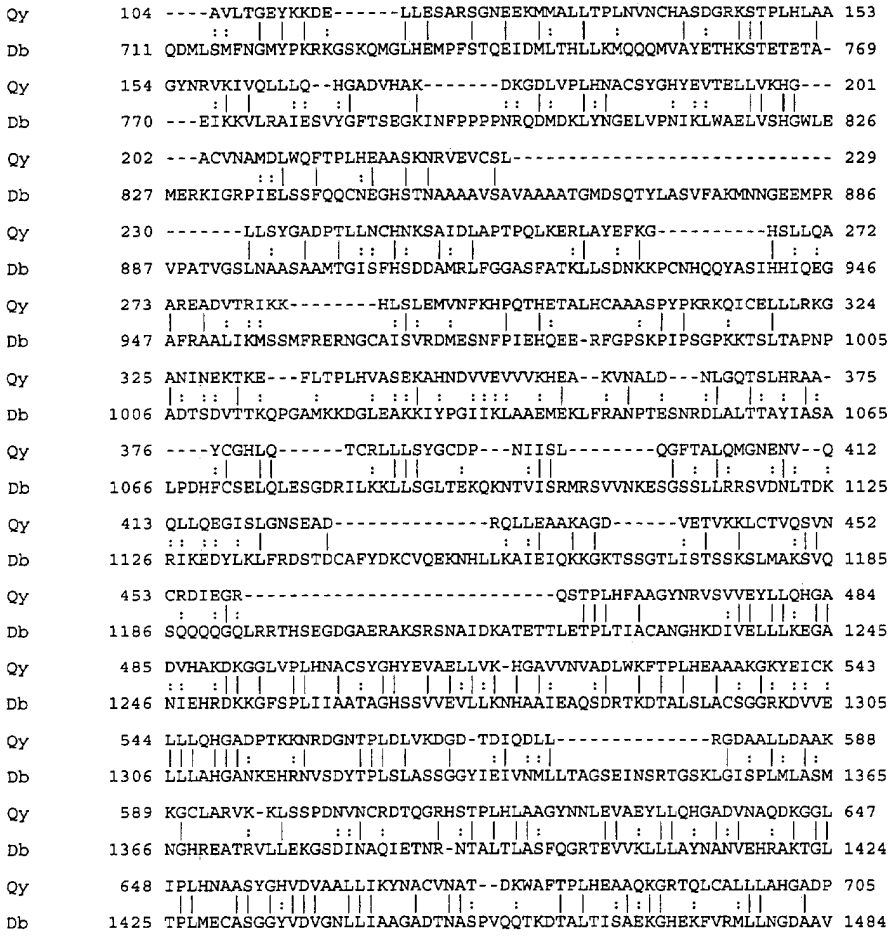




RESULT 29
T46507
hypothetical protein DKFZp586M2121.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46507
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46507
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-557 <AAA>
A;Cross-references: EMBL:AL137735
A;Experimental source: adult uterus; clone DKFZp586M2121
C;Genetics:
A;Note: DKFZp586M2121.1

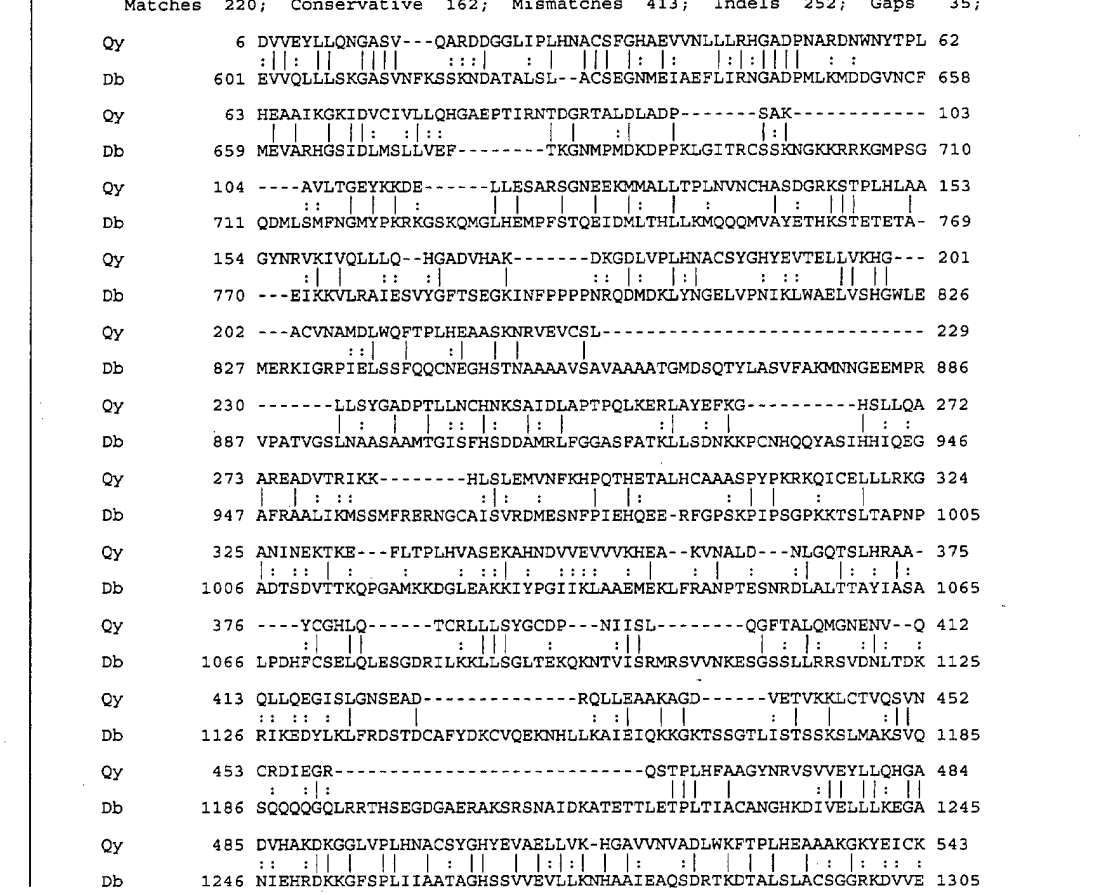
Query Match 6.1%; Score 350; DB 2; Length 557;
Best Local Similarity 24.3%; Pred. No. 4.1e-14;
Matches 133; Conservative 72; Mismatches 190; Indels 152; Gaps 15;

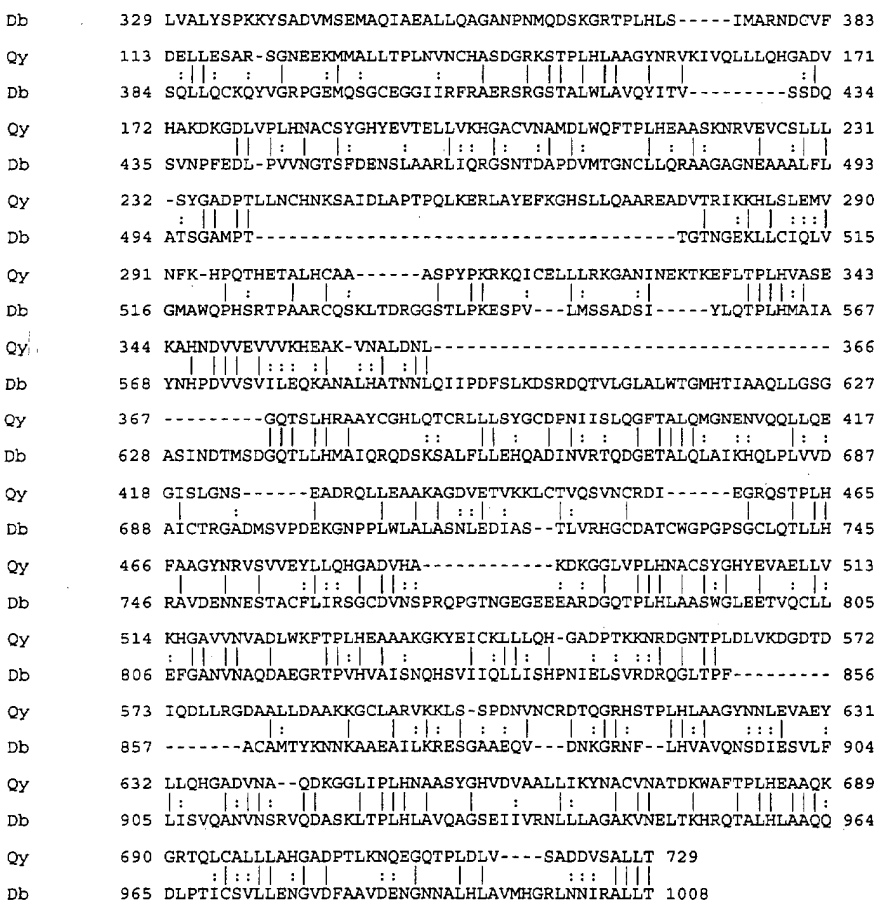




RESULT 27
T24157
hypothetical protein R11A8.7a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T24157; T24177
R;Bardill, S.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19846
A;Accession: T24157
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2606 <WIL>
A;Cross-references: EMBL:Z70686; PIDN:CAA94615.1; GSPDB:GN00022; CESP:R11A8.7a
A;Experimental source: clone R10H10
R;Cummings, P.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19849
A;Accession: T24177
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2606 <WI2>
A;Cross-references: EMBL:Z70310; PIDN:CAA94370.1; GSPDB:GN00022; CESP:R11A8.7a
A;Experimental source: clone R11A8
C;Genetics:
A;Gene: CESP:R11A8.7a
A;Map position: 4
A;Introns: 24/3; 80/3; 137/2; 177/1; 218/3; 348/3; 507/2; 709/3; 753/2; 924/1; 1020/3; 1

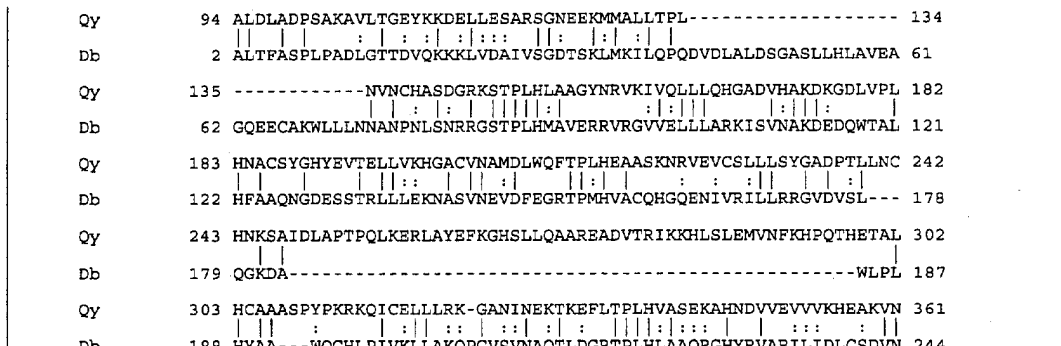
Query Match 6.5%; Score 372.5; DB 2; Length 2606;
Best Local Similarity 21.0%; Pred. No. 1.8e-14;
Matches 220; Conservative 162; Mismatches 413; Indels 252; Gaps 35;





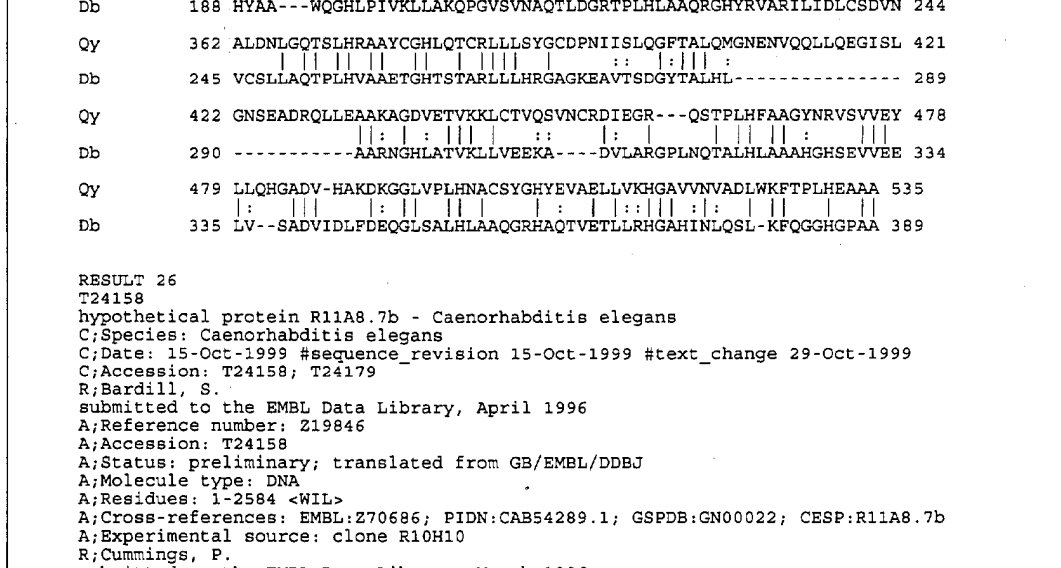
RESULT 25
T46445
hypothetical protein DKFZp434B2328.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46445
R;Bloecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23032
A;Accession: T46445
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-397 <AAA>
A;Cross-references: EMBL:AL137448
A;Experimental source: adult testis; clone DKFZp434B2328
C;Genetics:
A;Note: DKFZp434B2328.1

Query Match 6.7%; Score 388; DB 2; Length 397;
Best Local Similarity 27.4%; Pred. No. 9.8e-17;
Matches 131; Conservative 58; Mismatches 163; Indels 126; Gaps 11;



RESULT 26
T24158
hypothetical protein R11A8.7b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T24158; T24179
R;Bardill, S.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z19846
A;Accession: T24158
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2584 <WIL>
A;Cross-references: EMBL:Z70686; PIDN:CAB54289.1; GSPDB:GN00022; CESP:R11A8.7b
A;Experimental source: clone R10H10
R;Cummings, P.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19849
A;Accession: T24179
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2584 <WI2>
A;Cross-references: EMBL:Z70310; PIDN:CAB54294.1; GSPDB:GN00022; CESP:R11A8.7b
A;Experimental source: clone R11A8
C;Genetics:
A;Gene: CESP:R11A8.7b
A;Map position: 4
A;Introns: 24/3; 80/3; 137/2; 177/1; 218/3; 348/3; 507/2; 709/3; 753/2; 924/1; 1020/3; 1

Query Match 6.5%; Score 372.5; DB 2; Length 2584;
Best Local Similarity 21.0%; Pred. No. 1.8e-14;
Matches 220; Conservative 162; Mismatches 413; Indels 252; Gaps 35;



```
  Query Match             7.5%;  Score 431.5;  DB 2;  Length 426;
  Best Local Similarity   28.9%;  Pred. No. 2e-19;
  Matches  149;  Conservative   63;  Mismatches  182;  Indels  121;  Gaps   12;

Qy         11 LLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPN-ARDNWNYTPLHEAAIKG 69
              ||  |  |   |  |   |  | : |: |:|  ||  ||: | ||  :  | |  ||
Db         23 LLAAGVGVDICDRDGTTALMFAANLGYTEIVRSLLDGGANVNLARKRYGLTALMLAASAN 82

Qy         70 KIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMA 129
              ::|:  :|:  ||    | ||||      |: :|  || |
Db         83 QVDIVQLLISRGAAVNATNEDGSTA-------------------LMAAALKGNVEVARV 122

Qy        130 LLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYG 189
              ||     : : :|    | | ||    :  :|||:|  ||| : :|:     |  |   |
Db        123 LLA-AGADVNITDKDDDTALKLAVKRGQAAVVQLILPSGADANCEDEEGETLLMLAADSG 181

Qy        190 HYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAID 249
              | :| ::|:  |  ||  :   | |  | :     :  :||  |||
Db        182 HGDVVQVLLAAGVDVNEQNQDGGTALLAAVAAGNGAIAKILLDRGAD------------- 228

Qy        250 LAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASP 309
                                                  || :  |  |:||| |
Db        229 ---------------------------------------VNHQ-DQDGESALHLATVEG 247

Qy        310 YPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQT 369
              |     : ::|| :|||   | |   ||| ||: : |: :|| ::|: | |:  ||||:|
Db        248 Y---VDVVQVLLNQGANTQIKNKLGDTPLLVAALQGHDQIVETLLKYGANVHG-DNLGET 303

Qy        370 SLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN-----VQQLLQEGISLGNS 424
              |  ||  ||  | |:|| || : || :  | |||    |      :| || :|
Db        304 PLTLAASQGHTATVRILLDYGANANIPASDGKTALIKATERNHPGVIQLLLAKG------ 357

Qy        425 EADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAA--GYNRVSVVEYLLQH 482
                                          :|| :|  |  :| | :||  |||:  ||: ||:
Db        358 ------------------------ANVNYQDSVG--ATALIWAASGGYNK--VVQILLEG 389

Qy        483 GADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA 517
              ||| : |::||   |  |   |  : ::| : ||
Db        390 GADTNLKNRGGYTALMIAEFNGFRSIVQILKQAGA 424


RESULT 23
T25412
hypothetical protein T28D6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25412
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20031
A;Accession: T25412
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-991 <WIL>
A;Cross-references: EMBL:Z81134; PIDN:CAB03446.1; GSPDB:GN00021; CESP:T28D6.4
A;Experimental source: clone T28D6
C;Genetics:
A;Gene: CESP:T28D6.4
A;Map position: 3
A;Introns: 54/3; 255/3; 422/3; 584/1; 740/3; 892/3

  Query Match             7.4%;  Score 426;  DB 2;  Length 991;
  Best Local Similarity   25.6%;  Pred. No. 1.7e-18;
  Matches  185;  Conservative  100;  Mismatches  272;  Indels  166;  Gaps   26;

Qy        116 LESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGAD-VHAK 174
              |:|  :||          | : ||  |   :| |  :||         : ::  | |:|
Db        361 LQSVGAGN----------LVLRCHVID-LPTTALLKSAGAR-------MEENNYDVVYAC 402

Qy        175 DKGDL------------VP------LHNACSYGHYEVTELLVKH------GACVNAMDLW 210
                ||:             ||       |  | : |: |:      |  |       :|
Db        403 SVGDVSYLDAILREMKPVPMDICFGLLTAAARGNIEMCRFLADHFPHLVASSCCG----- 457

Qy        211 QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL 270
              ::  |  ||   ::|:  |||  | |           :|                 |||
Db        458 EWNALRSAACYGQLEILHLLLEKGVD----------VDEC-------------GHS-- 490

Qy        271 QAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEK 330
                                              |||  ||   :  :   :|||: || :: :
Db        491 --------------------------DRTALRAAA---WSGQLAAVQLLLQTGAEVDRR 520

Qy        331 TKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ-----TCRL 385
                |  |  |:   | ||| |:::: | |||:|  | |:||      |     | :|
Db        521 DSEQRTALMAAAFMCHKDVVSVIIQYGADVNAIDKSGATALHLNLSNGSKQEENSETTKL 580

Qy        386 LLSYGCDPNIISLQGFTALQM----GNENVQQLLQEG--ISLGNSEADRQLLEAAKAGDV 439
              || :  |  |     |  || :    |:  :| :|::   : : ::      |: ||  | :
Db        581 LLEHNADCKIEDANGRVALHLAAYHGDPCLQLILEKHPIVDVMDNMGQTPLMLAASQGQL 640

Qy        440 ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN 499
              | |: |       :   |:    | | :||       |||:|   |:|   ||  | | ||
Db        641 EAVQFLTKTAKADVDAIDNNGRTALQWAAINGHPEVVEFLCSVGSDEGHKDNDGAVALHY 700

Qy        500 ACSYGHYEVAELLVKH---GAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKN 556
              | :: | |    |||      |   | :   |:  |||  |  :: : |:       : :
Db        701 AVTHDHVE----LVKHLTNQHTVEAKDRYGQHPMMIAAANGNLKVLEYLMGISDIINQPD 756

Qy        557 RDGNTPLDLVK-DGDTDIQDLLRGDAALLDAAKK-------------GCLARVKKLSSPD 602
                :| | | |     | | ||:    : | :::                 |  | ::
Db        757 HEGRTALSLASMAAHTSIIDLI--SPRITDWSQRDHDGTPLAHTLILSNLIFVAEVYLGH 814

Qy        603 NVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQ-HGADVNAQDKGGLIPLHNAASYGHVDV 661
                |  |:     | :|: |  |::  |: | :  | :  | |:||    |  |     || :
Db        815 GANSSDSDAHGRTCVHVVASTNDVAAAKMLKRIGGVNFEAVDRGGRTALMTAVWAKHVAI 874

Qy        662 AALLIKYNAC---VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718
              :  |:  ::|     | ||:   | |  ||| |    |  |||  ||:| :|:  |:| :|:
Db        875 SVFLL--DSCGVDPNRTDRQGATALSIAAQIGDRDLVNLLLKFGAEPEIKDSMGRTAMDV 932

Qy        719 VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSS------LSAA 772
                    |:::   |  :|        |  |      ::   || |:||        | |
Db        933 -------AMISGNEPIVSL--------LQSVNGSSGSSGIGSSVPNSPLDASKQRILKAT 977

Qy        773 SSL 775
              |||
Db        978 SSL 980


RESULT 24
T00253
gene Ankhzn protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 05-Nov-1999
C;Accession: T00253
R;Ito, K.; Miyashita, A.; Ishii, N.; Kuriyama, H.; Tominaga, K.; Saitoh, H.; Maruyama, H.
submitted to the EMBL Data Library, February 1998
A;Description: CDNA cloning of mouse Ankhzn.
A;Reference number: Z14132
A;Accession: T00253
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1184 <ITO>
A;Cross-references: EMBL:AB011370; NID:d1179960; PIDN:BAA24980.1; PID:d1025905
C;Genetics:
A;Note: Ankhzn

  Query Match             7.0%;  Score 405.5;  DB 2;  Length 1184;
  Best Local Similarity   24.2%;  Pred. No. 4.4e-17;
  Matches  199;  Conservative  112;  Mismatches  330;  Indels  183;  Gaps   26;

Qy          5 KDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARD-NWNYTPLH 63
              : :   |: : | |    |  |   ||      |:::||  ||       ||||
Db        269 ESIATTLVSHKADVDMVDKNGWSLLHKGIQRGDLFASTFLIKNGALVNAATAGAQETPLH 328

Qy         64 EAAI----KGKIDV-------CIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKK 112
                |:    |   ||          ||| || | :::: ||| | |:      :   :
```

```
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-031-485-36

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 9.3e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       5058 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 4999

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       4998 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 4939

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db       4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 4879

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db       4878 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 4819

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db       4818 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4759

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db       4758 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4699

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db       4698 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4640

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db       4639 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       4219 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 4175

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

Qy       2582 A 2582
Db       3874 A 3874


RESULT 25
US-08-847-429A-35
; Sequence 35, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
```

```
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-35

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 9.3e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db        178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 237

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db        238 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 297

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db        298 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 357

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db        358 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 417

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db        418 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 477

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db        478 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 537

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db        538 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 596

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db        597 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db        657 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db        717 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 776

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db        777 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 836

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db        837 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 896

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db        897 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 956

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db        957 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db        1017 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 1061

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db        1062 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1121

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db        1122 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1181

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db        1182 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db        1242 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1301

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db        1302 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1361

Qy       2582 A 2582
              |
Db       1362 A 1362
```

```
RESULT 26
US-08-847-429A-36/c
; Sequence 36, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  36:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-36

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 9.3e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       5058 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 4999

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       4998 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 4939

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db       4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 4879

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db       4878 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 4819

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db       4818 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4759

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db       4758 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4699

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db       4698 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4640

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db       4639 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       4219 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 4175

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
```

```
Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

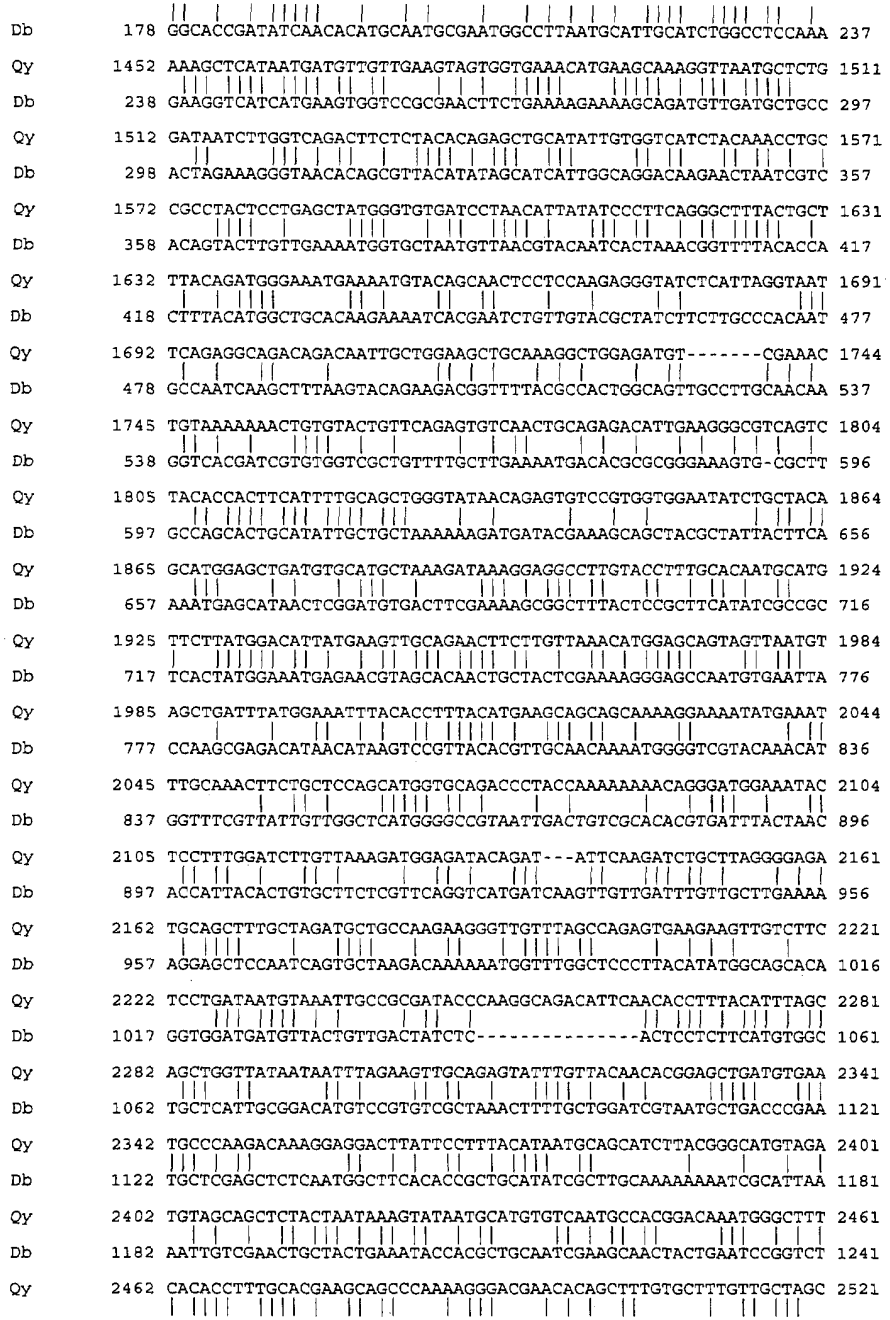
Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

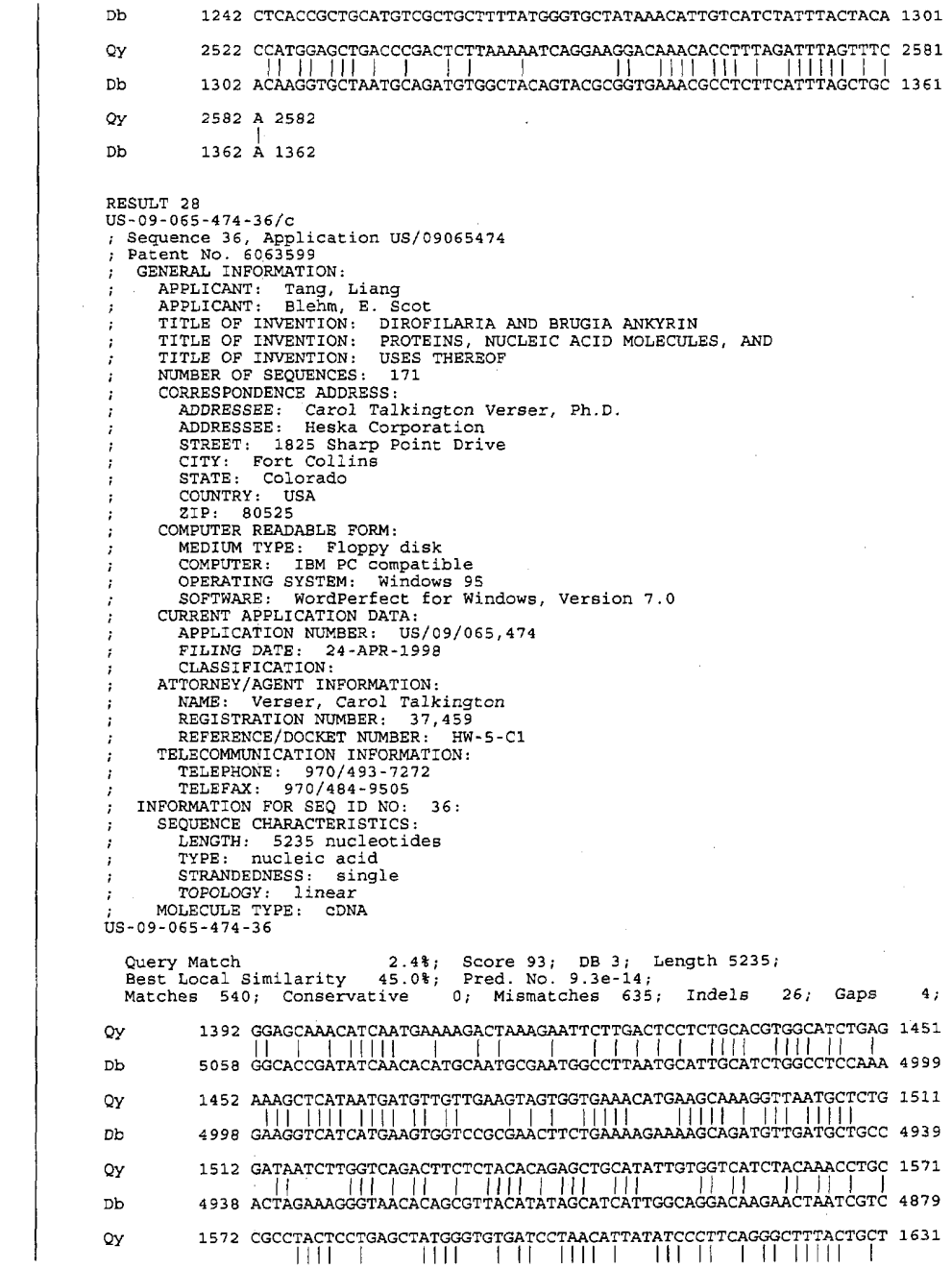
Qy       2582 A 2582
Db       3874 A 3874
```

```
RESULT 27
US-09-065-474-35
; Sequence 35, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-35

  Query Match             2.4%;  Score 93;  DB 3;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 9.3e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

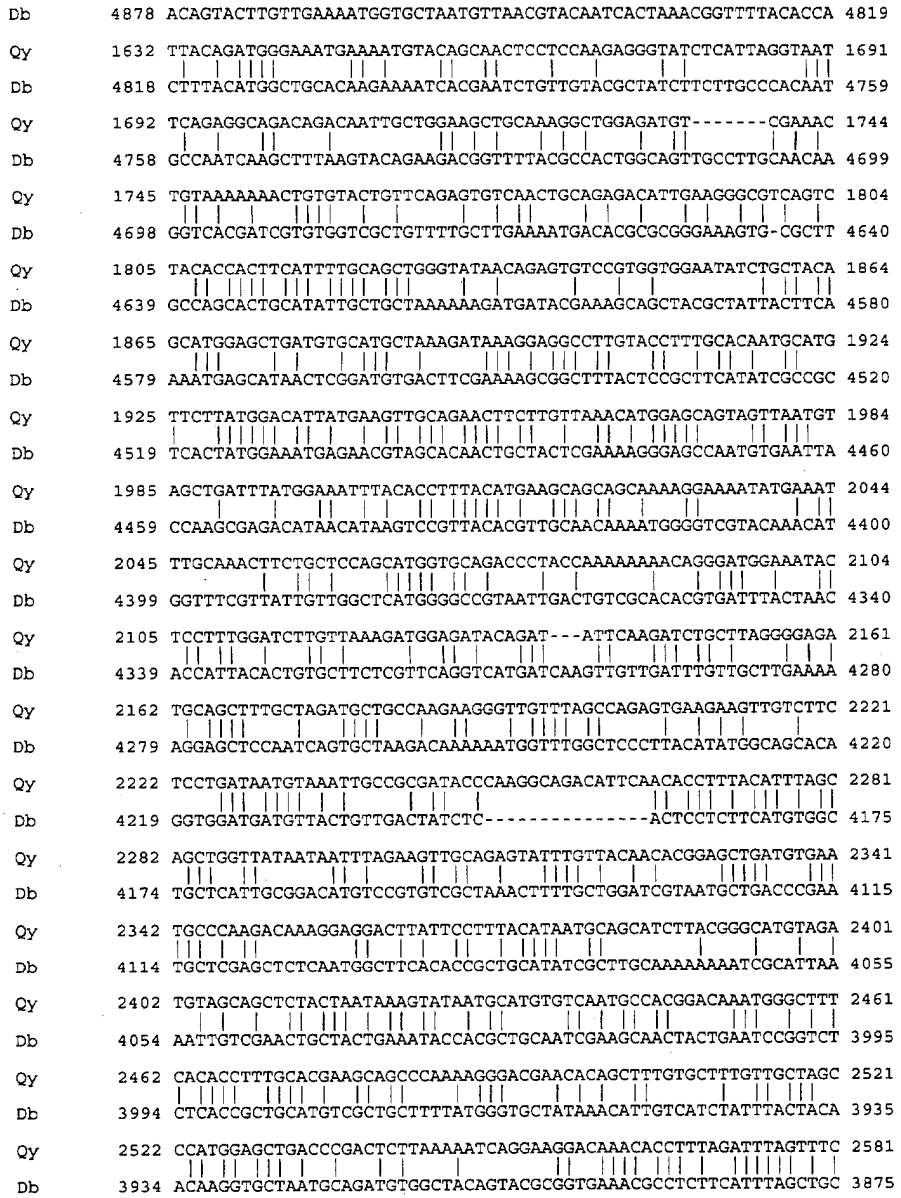
Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
```





RESULT 28
US-09-065-474-36/c
; Sequence 36, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-065-474-36

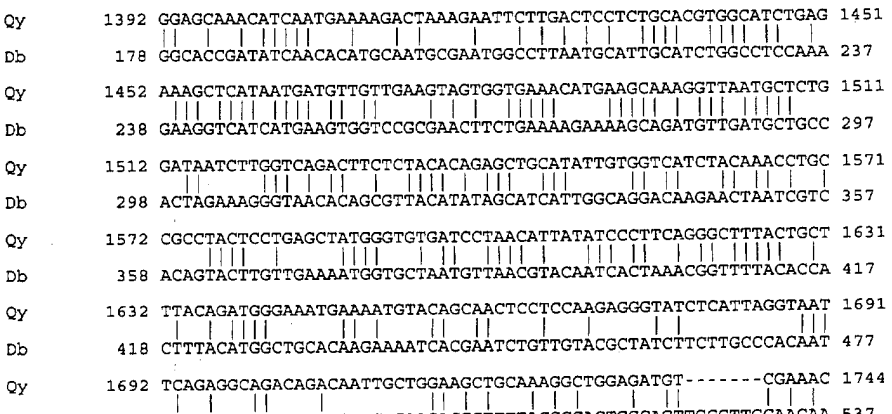
Query Match 2.4%; Score 93; DB 3; Length 5235;
Best Local Similarity 45.0%; Pred. No. 9.3e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

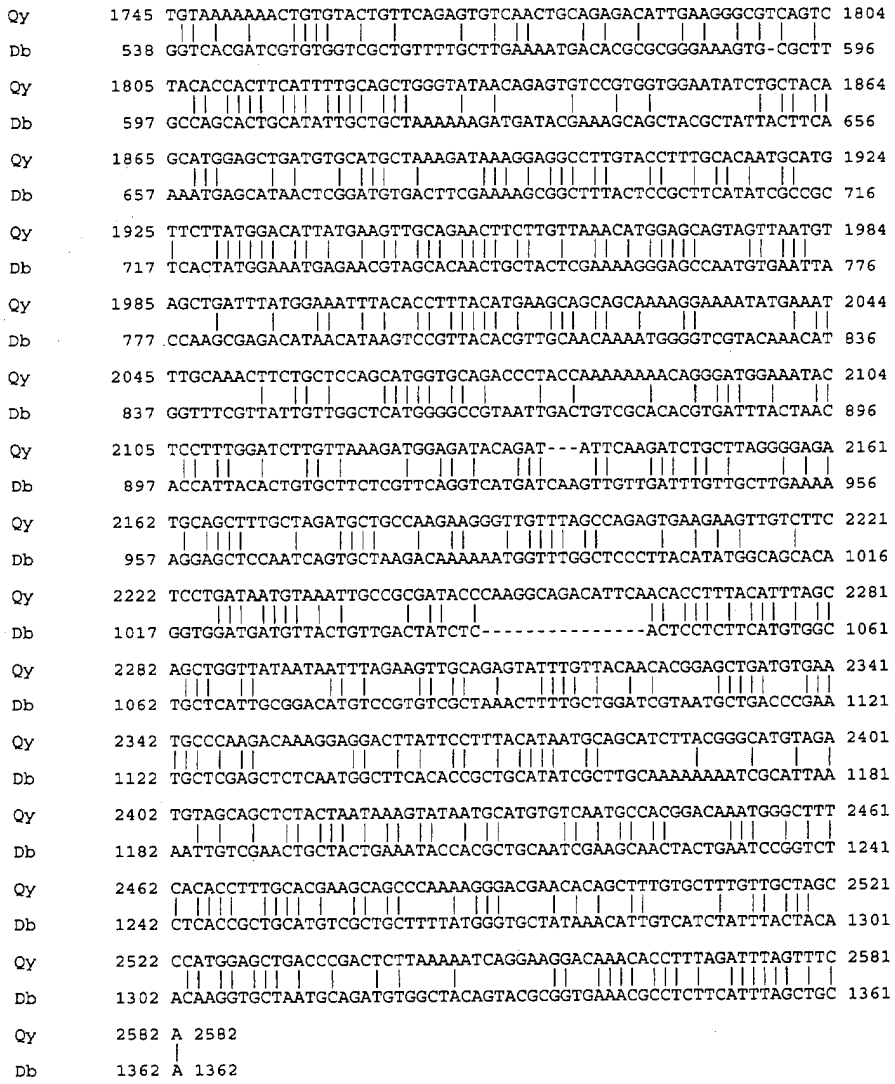


RESULT 29
US-09-557-034-35

; Sequence 35, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-557-034-35

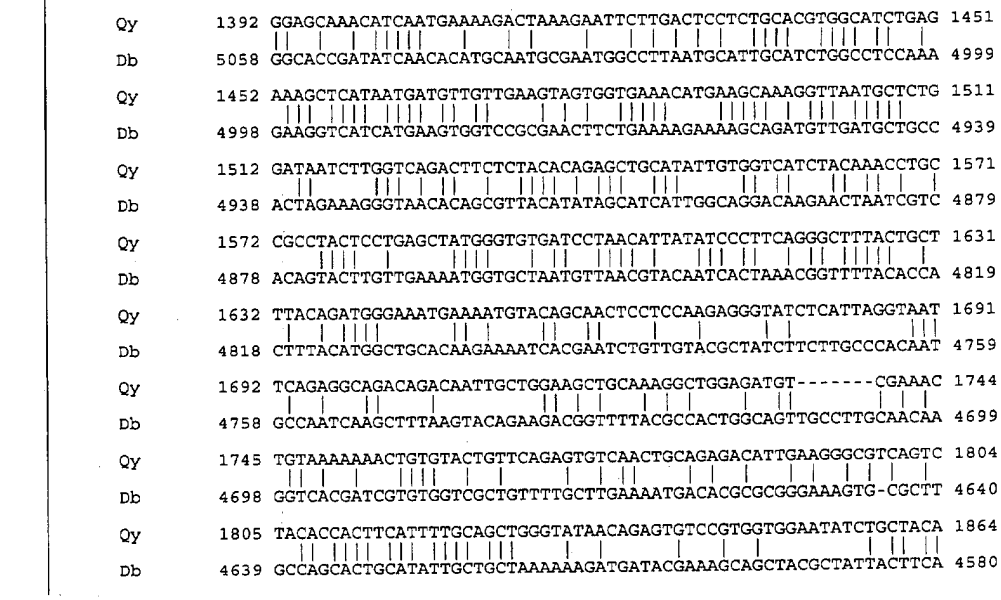
Query Match 2.4%; Score 93; DB 4; Length 5235;
Best Local Similarity 45.0%; Pred. No. 9.3e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;





```
RESULT 30
US-09-557-034-36/c
; Sequence 36, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Liang
;                   Blehm, E. Scot
;        TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                            PROTEINS, NUCLEIC ACID MOLECULES, AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 171
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Carol Talkington Verser, Ph.D.
;                        Heska Corporation
;             STREET: 1825 Sharp Point Drive
;             CITY: Fort Collins
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80525
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: WordPerfect for Windows, Version 7.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,034
;             FILING DATE: 21-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/065,474
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Verser, Carol Talkington
;             REGISTRATION NUMBER: 37,459
;             REFERENCE/DOCKET NUMBER: HW-5-C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 970/493-7272
;             TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 36:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5235 nucleotides
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-557-034-36

  Query Match             2.4%;  Score 93;  DB 4;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 9.3e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;
```



```
Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       4219 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 4175

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

Qy       2582 A 2582
              |
Db       3874 A 3874
```

```
RESULT 31
US-09-031-485-32
; Sequence 32, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  51..5285
US-09-031-485-32

  Query Match             2.4%;  Score 93;  DB 1;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 9.5e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db        228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db        288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db        348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db        408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db        468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

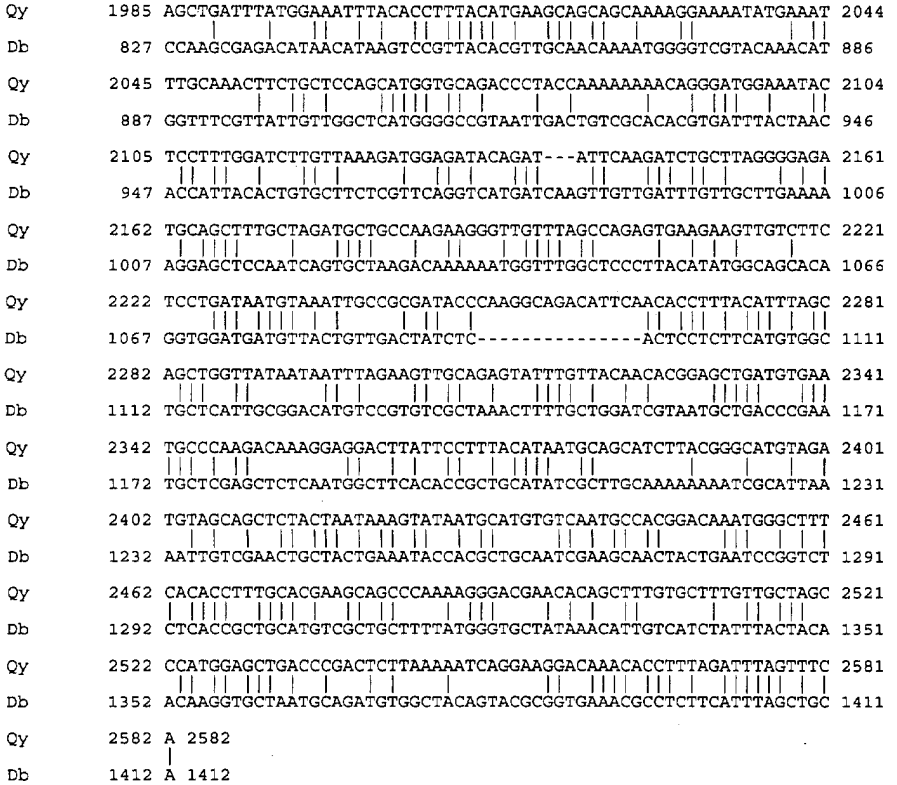
Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db        528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db        588 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db        647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

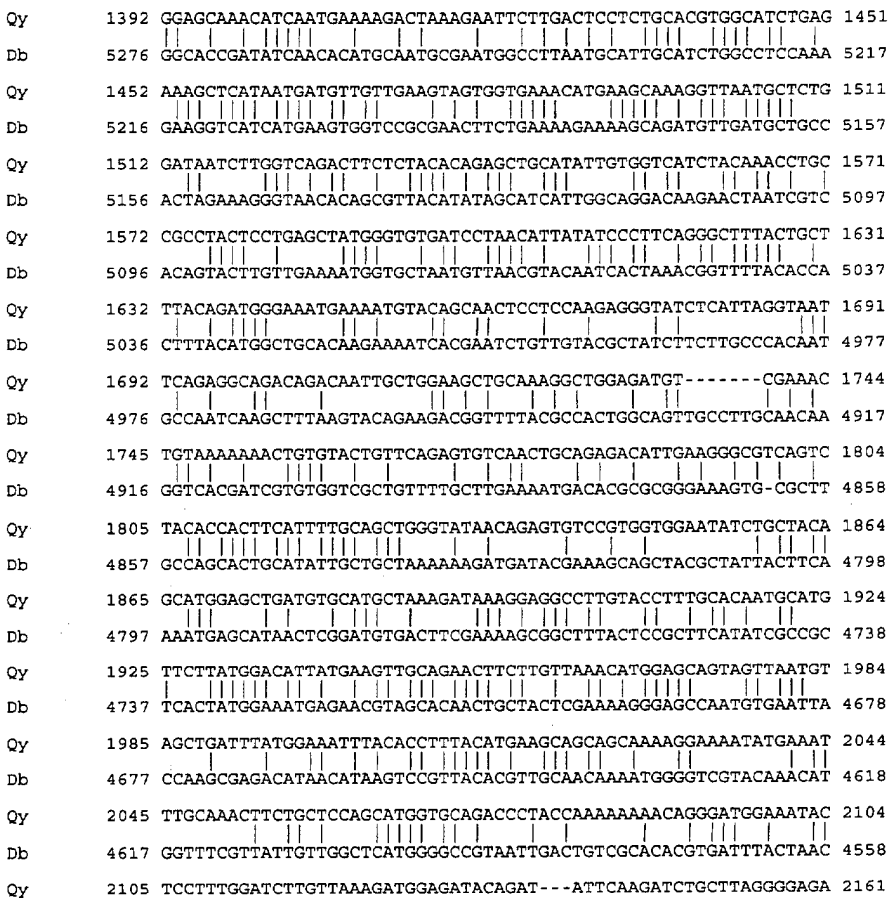
Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db        707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

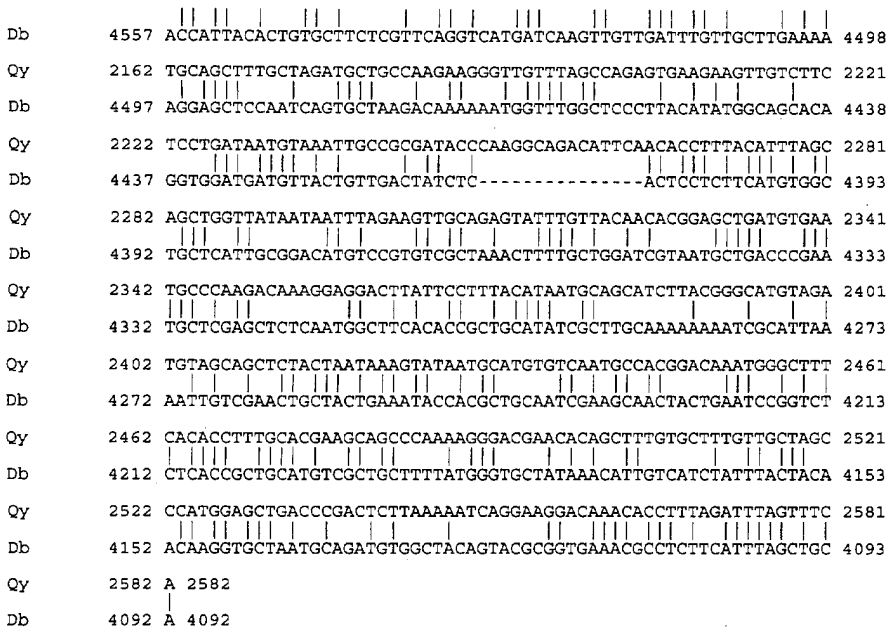
Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db        767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826
```



RESULT 32
US-09-031-485-34/c
; Sequence 34, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-031-485-34

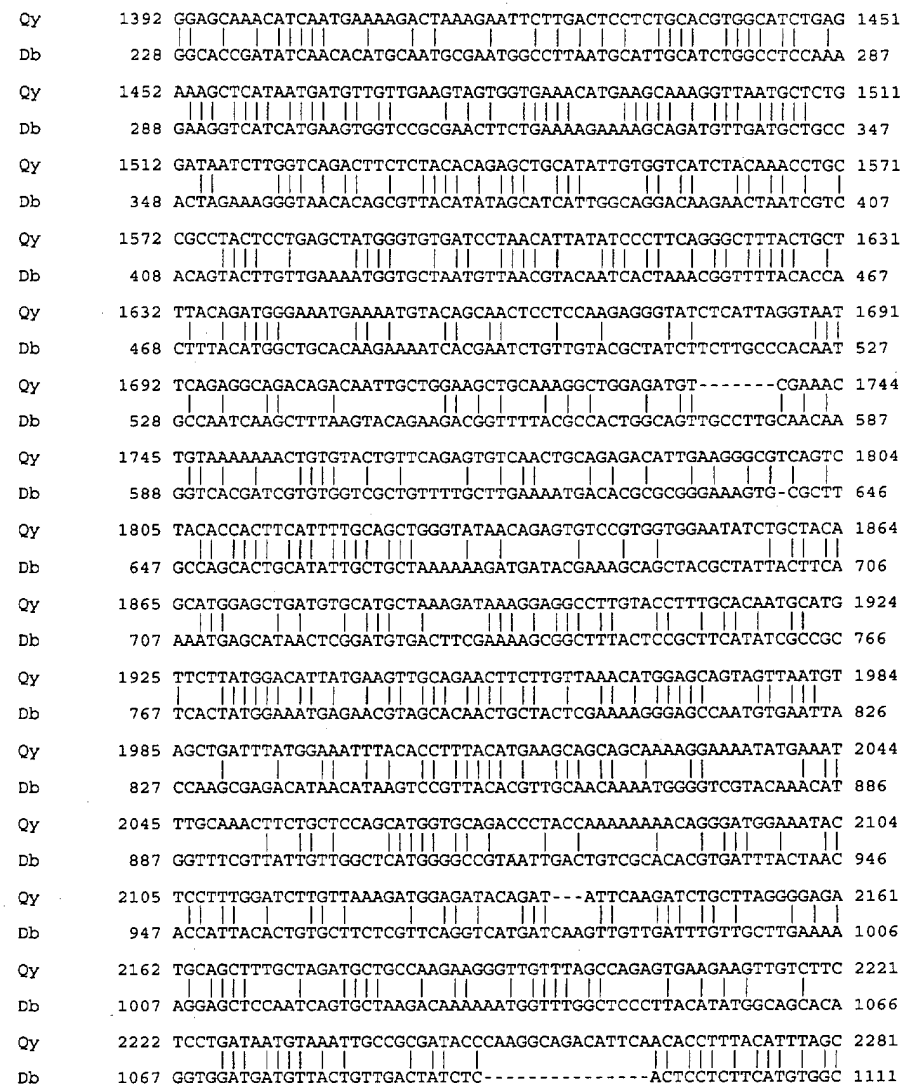
Query Match 2.4%; Score 93; DB 1; Length 5503;
Best Local Similarity 45.0%; Pred. No. 9.5e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;





RESULT 33
US-08-847-429A-32
; Sequence 32, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..5285
US-08-847-429A-32

Query Match 2.4%; Score 93; DB 1; Length 5503;
Best Local Similarity 45.0%; Pred. No. 9.5e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;





```
Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       1112 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1171

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       1172 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1231

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       1232 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1291

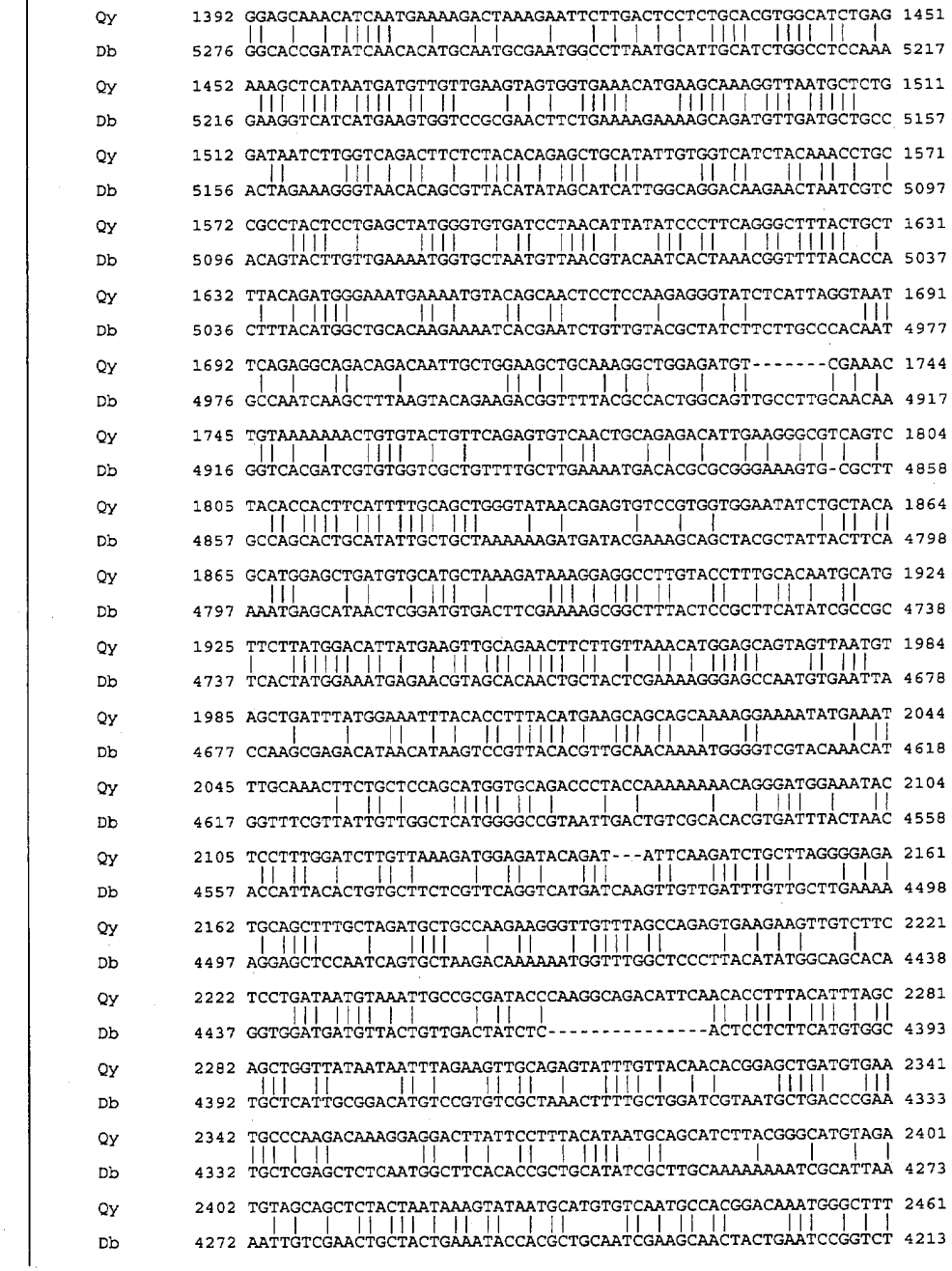
Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       1292 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1351

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       1352 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1411

Qy       2582 A 2582
              |
Db       1412 A 1412
```

RESULT 34
US-08-847-429A-34/c
; Sequence 34, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-847-429A-34

Query Match 2.4%; Score 93; DB 1; Length 5503;
Best Local Similarity 45.0%; Pred. No. 9.5e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;



```
Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       5276 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 5217

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       5216 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 5157

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db       5156 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 5097

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db       5096 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 5037

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db       5036 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4977

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db       4976 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4917

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db       4916 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4858

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db       4857 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4798

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db       4797 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4738

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db       4737 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4678

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db       4677 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4618

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db       4617 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4558

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db       4557 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4498

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       4497 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4438

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       4437 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 4393

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213
```

```
Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2582 A 2582
              |
Db       4092 A 4092
```

RESULT 35
US-09-065-474-32
; Sequence 32, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..5285
US-09-065-474-32

Query Match 2.4%; Score 93; DB 3; Length 5503;
Best Local Similarity 45.0%; Pred. No. 9.5e-14;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

```
Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db        228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db        288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db        348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db        408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db        468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db        528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db        588 GGTCACGATCGTGTGGTCGCTGTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db        647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db        707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db        767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db        827 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 886

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db        887 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 946

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db        947 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 1006

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       1007 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1066

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       1067 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 1111

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       1112 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1171

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       1172 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1231

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       1232 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1291

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       1292 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1351

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       1352 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1411

Qy       2582 A 2582
              |
```

```
Db       1412 A 1412


RESULT 36
US-09-065-474-34/c
; Sequence 34, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-34

  Query Match             2.4%;  Score 93;  DB 3;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 9.5e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       5276 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 5217

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       5216 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 5157

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db       5156 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 5097

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db       5096 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 5037

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db       5036 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4977

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744

Db       4976 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4917

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db       4916 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4858

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
Db       4857 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4798

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db       4797 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4738

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db       4737 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4678

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db       4677 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4618

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db       4617 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4558

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db       4557 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4498

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       4497 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4438

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       4437 GGTGGATGATGTTACTGTTGACTATCTC----------------ACTCCTCTTCATGTGGC 4393

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2582 A 2582
Db       4092 A 4092


RESULT 37
US-09-557-034-32
; Sequence 32, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Liang
;                   Blehm, E. Scot
;        TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                            PROTEINS, NUCLEIC ACID MOLECULES, AND
;                            USES THEREOF
;
```

```
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 32:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 5503 nucleotides
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;                  (ix)   FEATURES:
;       SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-557-034-32

  Query Match             2.4%;  Score 93;  DB 4;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 9.5e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db        228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db        288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db        348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db        408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
Db        468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
Db        528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
Db        588 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
```

```
Db        647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
Db        707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
Db        767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
Db        827 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 886

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
Db        887 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 946

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
Db        947 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 1006

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db        1007 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1066

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db        1067 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 1111

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db        1112 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1171

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db        1172 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1231

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db        1232 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1291

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db        1292 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1351

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db        1352 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1411

Qy       2582 A 2582
Db        1412 A 1412


RESULT 38
US-09-557-034-34/c
; Sequence 34, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
```

```
;           ZIP: 80525
;       COMPUTER READABLE FORM:
;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: Windows 95
;           SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/557,034
;           FILING DATE: 21-Apr-2000
;           CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;           APPLICATION NUMBER: 09/065,474
;           FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;           NAME: Verser, Carol Talkington
;           REGISTRATION NUMBER: 37,459
;           REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: 970/493-7272
;           TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 34:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 5503 nucleotides
;           TYPE: nucleic acid
;           STRANDEDNESS: single
;           TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-557-034-34

  Query Match             2.4%;  Score 93;  DB 4;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 9.5e-14;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
             ||  |  | |||||  |  | |  |  | | | | |||| |||| || |
Db       5276 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 5217

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
              ||| |||| |||| || ||    | | | |||||    ||||| | ||| |||||
Db       5216 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 5157

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
               ||     ||| | ||  |  |||| | ||| |||     || ||  || || | |
Db       5156 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 5097

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
              |||| |    ||||  | ||  |||| |   ||| ||  | || |||||  |
Db       5096 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 5037

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
              |  | ||||     || |   ||  ||   | |   | |           |||
Db       5036 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4977

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1744
              |  |  ||    |          || | |  | | |   |  ||        | | |
Db       4976 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4917

Qy       1745 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1804
              || |  |  |||| |  |    |  | ||   | |  |  | |  | | | |
Db       4916 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4858

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
              || |||| ||| |||| |||    |  |     |  |  |          | || ||
Db       4857 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4798

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
              |||  | |  | |||  |    ||| | ||| ||  ||  | || |  ||
Db       4797 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4738

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
              |  |||||| || |  | || ||| |||| ||  |  || | |||||    || |||
Db       4737 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4678

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
              |   |  ||  | |  || |||||  |  ||| ||  |  ||    |  ||
Db       4677 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4618

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2104
                 |  || |   ||||| || |  |  |  |  |  |  | ||| |  ||
Db       4617 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4558

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2161
              ||  ||  |  || |    | || |  |||    ||  ||| || |   | | |
Db       4557 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4498

Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
              | ||||    |  |||  |  ||  | |||| ||     |  | | |  |
Db       4497 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4438

Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
                 ||| |||| | |    | ||  |                || ||| | ||| | ||
Db       4437 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 4393

Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
              |||  ||    || |  || ||  | |||| | | |    ||||   |||
Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
              |||  | ||           || | |  ||  | |||| ||       |   | | |
Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
              | ||  | ||  | || | | || ||  | ||      || | || ||   ||| | | |
Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213

Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
              | |||| |||| |  || |||    | |||  | | |  ||      | || |||
Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
              || || ||| |  |  | |  |     ||  |||| ||| |  |||||| | |
Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2582 A 2582
              |
Db       4092 A 4092


RESULT 39
US-09-065-474-138
; Sequence 138, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
```

```
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Verser, Carol Talkington
;       REGISTRATION NUMBER:  37,459
;       REFERENCE/DOCKET NUMBER:  HW-5-C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  970/493-7272
;       TELEFAX:  970/484-9505
;   INFORMATION FOR SEQ ID NO:  138:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1056 nucleotides
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1056
US-09-065-474-138

  Query Match             2.2%;  Score 83.4;  DB 3;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.2e-11;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2259 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2318
             ||   ||||||| | ||| |||| ||   ||| |||||     | || |||   ||    |
Db        427 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 486

Qy       2319 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAT 2378
              || || | || |||  ||  |||||  | |||    ||    | ||||||||||| |
Db        487 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 546

Qy       2379 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2438
              || ||      |    ||  | ||| | ||         || |  | ||||| ||| ||
Db        547 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 606

Qy       2439 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2498
              |||||      ||||  |  || |||||  | ||    || |||||  |||| |
Db        607 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 666

Qy       2499 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2558
              |  |  ||| || ||  | |  |||||| |      | |     | | || | ||
Db        667 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 726

Qy       2559 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2591
                |||||  |  |||||  | ||  ||| |||
Db        727 TTGACACCAATGCATTTATGTGCACAGGAGGAT 759


RESULT 40
US-09-065-474-140/c
; Sequence 140, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1056 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-140

  Query Match             2.2%;  Score 83.4;  DB 3;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.2e-11;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2259 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2318
             ||   ||||||| | ||| |||| ||   ||| |||||     | || |||   ||    |
Db        630 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 571

Qy       2319 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAT 2378
              || || | || |||  ||  |||||  | |||    ||    | ||||||||||| |
Db        570 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 511

Qy       2379 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2438
              || ||      |    ||  | ||| | ||         || |  | ||||| ||| ||
Db        510 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 451

Qy       2439 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2498
              |||||      ||||  |  || |||||  | ||    || |||||  |||| |
Db        450 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 391

Qy       2499 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2558
              |  |  ||| || ||  | |  |||||| |      | |     | | || | ||
Db        390 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 331

Qy       2559 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2591
                |||||  |  |||||  | ||  ||| |||
Db        330 TTGACACCAATGCATTTATGTGCACAGGAGGAT 298


RESULT 41
US-09-557-034-138
; Sequence 138, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
```

```
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/557,034
;              FILING DATE: 21-Apr-2000
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/065,474
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Verser, Carol Talkington
;              REGISTRATION NUMBER: 37,459
;              REFERENCE/DOCKET NUMBER: HW-5-C1
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 970/493-7272
;              TELEFAX: 970/484-9505
;    INFORMATION FOR SEQ ID NO: 138:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1056 nucleotides
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;                        (ix)   FEATURES
;         SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-557-034-138

  Query Match             2.2%;  Score 83.4;  DB 4;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.2e-11;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2259 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2318
             ||   ||||||| | ||| |||| ||  ||| |||||     | || |||  ||   |
Db        427 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 486

Qy       2319 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAT 2378
             || || | || |||  ||   |||||   | |||    ||    | ||||||||||| |
Db        487 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 546

Qy       2379 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2438
             || ||      |    ||  | ||| | ||       ||  |  | ||||| ||| ||
Db        547 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 606

Qy       2439 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2498
             |||||       ||||   |  || |||||  | ||    || |||||  |||| |
Db        607 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 666

Qy       2499 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2558
               |   ||| || ||  | |  ||||||  |          | |      | | ||
Db        667 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 726

Qy       2559 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2591
                |||||  |  |||||  | ||  ||| |||
Db        727 TTGACACCAATGCATTTATGTGCACAGGAGGAT 759


RESULT 42
US-09-557-034-140/c
; Sequence 140, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;         APPLICANT: Tang, Liang
;                    Blehm, E. Scot
;         TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                             PROTEINS, NUCLEIC ACID MOLECULES, AND
;                             USES THEREOF
;         NUMBER OF SEQUENCES: 171
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Carol Talkington Verser, Ph.D.
;                         Heska Corporation
;              STREET: 1825 Sharp Point Drive
;              CITY: Fort Collins
;              STATE: Colorado
;              COUNTRY: USA
;              ZIP: 80525
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: Windows 95
;              SOFTWARE: WordPerfect for Windows, Version 7.0
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/557,034
;              FILING DATE: 21-Apr-2000
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/065,474
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Verser, Carol Talkington
;              REGISTRATION NUMBER: 37,459
;              REFERENCE/DOCKET NUMBER: HW-5-C1
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 970/493-7272
;              TELEFAX: 970/484-9505
;    INFORMATION FOR SEQ ID NO: 140:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1056 nucleotides
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         SEQUENCE DESCRIPTION: SEQ ID NO: 140:
US-09-557-034-140

  Query Match             2.2%;  Score 83.4;  DB 4;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.2e-11;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2259 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2318
             ||   ||||||| | ||| |||| ||  ||| |||||     | || |||  ||   |
Db        630 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 571

Qy       2319 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAT 2378
             || || | || |||  ||   |||||   | |||    ||    | ||||||||||| |
Db        570 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 511

Qy       2379 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2438
             || ||      |    ||  | ||| | ||       ||  |  | ||||| ||| ||
Db        510 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 451

Qy       2439 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2498
             |||||       ||||   |  || |||||  | ||    || |||||  |||| |
Db        450 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 391

Qy       2499 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2558
               |   ||| || ||  | |  ||||||  |          | |      | | ||
Db        390 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 331

Qy       2559 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2591
                |||||  |  |||||  | ||  ||| |||
Db        330 TTGACACCAATGCATTTATGTGCACAGGAGGAT 298


RESULT 43
US-09-031-485-25
; Sequence 25, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
```

```
;       CITY:  Fort Collins
;       STATE:  Colorado
;       COUNTRY:  USA
;       ZIP:  80525
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect for Windows, Version 7.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/031,485
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/847,429
;       FILING DATE:  24-APR-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Verser, Carol Talkington
;       REGISTRATION NUMBER:  37,459
;       REFERENCE/DOCKET NUMBER:  HW-5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  970/493-7272
;       TELEFAX:  970/484-9505
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  909 nucleotides
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-09-031-485-25

  Query Match             2.2%;  Score 82.6;  DB 1;  Length 909;
  Best Local Similarity   53.2%;  Pred. No. 1.7e-11;
  Matches  175;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy        402 ACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 461
              || || || || || || ||| |       | | ||| | |||     |||| | |
Db        325 ACGCCTCTTCATTTAGCTGCACGAGCAAACCAAACGGACATTGTTCGTGTTTTGGTGCGT 384

Qy        462 AATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 521
              ||||| ||| | || | ||   || | ||      | | ||| || || | |||||
Db        385 AATGGAGCACAGGTGGATGCTGCTGCTCGTGAACTACAAACTCCACTGCACATTGCATCA 444

Qy        522 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 581
              | |||||| || | ||  | |||| | |  | |||     | ||||   || ||||||
Db        445 CGTCTTGGTAATACCGACATCGTCATTTTGTTGCTGCAGGCTAATGCATCACCAAATGCT 504

Qy        582 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 641
              |   ||||||||||||| |||   ||||| |  | ||  |    || ||
Db        505 GCCACAAGAGATCTTTATACTCCTCTTCATATTGCTGCCAAGGAGGGGCAAGAGGAAGTG 564

Qy        642 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 701
              |  || |  | |||||| | ||   ||  | |  |  | | | ||    ||
Db        565 GCAGCAATATTGATGGATCATGGAACCGACAAGACACTGCTCACGAAAAAGGGTTTTACG 624

Qy        702 GCATTGGATTTAGCAGATCCATCTGCCAA 730
              | ||| |||||| | |   | || |||
Db        625 CCGTTGCATTTAGCTGCTAAGTATGGCAA 653


RESULT 44
US-09-031-485-26/c
; Sequence 26, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  909 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-031-485-26

  Query Match             2.2%;  Score 82.6;  DB 1;  Length 909;
  Best Local Similarity   53.2%;  Pred. No. 1.7e-11;
  Matches  175;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy        402 ACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 461
              || || || || || || ||| |       | | ||| | |||     |||| | |
Db        585 ACGCCTCTTCATTTAGCTGCACGAGCAAACCAAACGGACATTGTTCGTGTTTTGGTGCGT 526

Qy        462 AATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 521
              ||||| ||| | || | ||   || | ||      | | ||| || || | |||||
Db        525 AATGGAGCACAGGTGGATGCTGCTGCTCGTGAACTACAAACTCCACTGCACATTGCATCA 466

Qy        522 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 581
              | |||||  || | ||  | |||| | |  | |||     | ||||   || ||||||
Db        465 CGTCTTGGTAATACCGACATCGTCATTTTGTTGCTGCAGGCTAATGCATCACCAAATGCT 406

Qy        582 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 641
              |   ||||||||||||| |||   ||||| |  | ||  |    || ||
Db        405 GCCACAAGAGATCTTTATACTCCTCTTCATATTGCTGCCAAGGAGGGGCAAGAGGAAGTG 346

Qy        642 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 701
              |  || |  | |||||| | ||   ||  | |  |  | | | ||    ||
Db        345 GCAGCAATATTGATGGATCATGGAACCGACAAGACACTGCTCACGAAAAAGGGTTTTACG 286

Qy        702 GCATTGGATTTAGCAGATCCATCTGCCAA 730
              | ||| |||||| | |   | || |||
Db        285 CCGTTGCATTTAGCTGCTAAGTATGGCAA 257


RESULT 45
US-08-847-429A-25
; Sequence 25, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
```

```
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  909 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-25

  Query Match             2.2%;  Score 82.6;  DB 1;  Length 909;
  Best Local Similarity   53.2%;  Pred. No. 1.7e-11;
  Matches  175;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy        402 ACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 461
              || || || || || || ||| |        |  |  ||| | |||     |||| | |
Db        325 ACGCCTCTTCATTTAGCTGCACGAGCAAACCAAACGGACATTGTTCGTGTTTTGGTGCGT 384

Qy        462 AATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 521
              ||||| ||| | ||  | ||   || |  ||      |  | ||| || || | |||||
Db        385 AATGGAGCACAGGTGGATGCTGCTGCTCGTGAACTACAAACTCCACTGCACATTGCATCA 444

Qy        522 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 581
                | ||||| || | ||  | |||| | |  |  |||     |  ||||   || ||||||
Db        445 CGTCTTGGTAATACCGACATCGTCATTTTGTTGCTGCAGGCTAATGCATCACCAAATGCT 504

Qy        582 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 641
                  |       |||||||||||| |||   ||||| |   | ||  |    || ||
Db        505 GCCACAAGAGATCTTTATACTCCTCTTCATATTGCTGCCAAGGAGGGGCAAGAGGAAGTG 564

Qy        642 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 701
                | || |  | ||||| | ||    ||  | |  | | | | ||    ||
Db        565 GCAGCAATATTGATGGATCATGGAACCGACAAGACACTGCTCACGAAAAAGGGTTTTACG 624

Qy        702 GCATTGGATTTAGCAGATCCATCTGCCAA 730
              | |||  ||||||| | |   | || |||
Db        625 CCGTTGCATTTAGCTGCTAAGTATGGCAA 653


Search completed: March 30, 2004, 09:48:44
Job time : 302.12 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 19:26:40 ; Search time 14402.4 Seconds
(without alignments)
11484.026 Million cell updates/sec

Title: US-09-843-159B-2
Perfect score: 3816
Sequence: 1 cgcgctgctccgcccgccgc.....cctttgctgaaaaaaaaaa 3816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3816	100.0	3816	6	AR392873	AR392873 Sequence
2	3793	99.4	6028	9	AF309033	AF309033 Homo sapi
3	3793	99.4	6123	6	AX057579	AX057579 Sequence
4	3791.4	99.4	4275	6	AR361486	AR361486 Sequence
5	3791.4	99.4	4406	6	AX062247	AX062247 Sequence
6	3791.4	99.4	4992	6	AX062275	AX062275 Sequence
7	3791.4	99.4	5002	6	AX062273	AX062273 Sequence
8	3791.4	99.4	6189	9	AF438201	AF438201 Homo sapi
9	3737	97.9	4295	9	AF305081	AF305081 Homo sapi
10	3724.8	97.6	3815	9	AF329696	AF329696 Homo sapi
11	3595.8	94.2	4127	6	AX062241	AX062241 Sequence
12	3594.8	94.2	4512	6	AR232161	AR232161 Sequence
13	3594.8	94.2	4512	6	AX076213	AX076213 Sequence
14	3585.8	94.0	5810	9	AF264912	AF264912 Homo sapi
15	3504.8	91.8	3508	6	AX062142	AX062142 Sequence
16	3499	91.7	3501	9	AF342982	AF342982 Homo sapi
17	3488.8	91.4	3498	6	AR232162	AR232162 Sequence
18	3488.8	91.4	3498	6	AX076214	AX076214 Sequence
19	3393.4	88.9	3797	6	AR392872	AR392872 Sequence
20	3387.6	88.8	3394	6	AR352002	AR352002 Sequence
21	3387.6	88.8	3394	6	AX133694	AX133694 Sequence
22	3341	87.6	4297	6	AR361485	AR361485 Sequence
23	3297	86.4	3400	6	AX029397	AX029397 Sequence
24	3297	86.4	3400	6	BD075726	BD075726 Potential
25	3039	79.6	4493	6	AR361484	AR361484 Sequence
26	2774.6	72.7	5005	6	AX052530	AX052530 Sequence
27	2522.6	66.1	3930	5	AY142107	AY142107 Gallus ga
28	2439.8	63.9	2971	6	AX062233	AX062233 Sequence
29	2439.8	63.9	3353	6	AX062229	AX062229 Sequence
30	2357.8	61.8	3799	6	AX062231	AX062231 Sequence
31	2122	55.6	3254	6	AX057580	AX057580 Sequence
32	1922.4	50.4	4130	9	HSM804989	AL833676 Homo sapi
33	1921.4	50.4	4054	10	BC057370	BC057370 Mus muscu
34	1906.4	50.0	3804	5	AY142108	AY142108 Gallus ga
35	1885	49.4	3984	6	AX062144	AX062144 Sequence
36	1885	49.4	4134	6	BD235110	BD235110 Tankyrase
37	1885	49.4	4134	6	AR274521	AR274521 Sequence
38	1885	49.4	4134	6	AR361487	AR361487 Sequence
39	1885	49.4	4134	6	AX467061	AX467061 Sequence
40	1885	49.4	4134	9	AF082556	AF082556 Homo sapi
41	1796.4	47.1	8901	9	AF082557	AF082557 Homo sapi
42	1568	41.1	1568	6	AX062239	AX062239 Sequence
43	1521.2	39.9	4491	6	BD235115	BD235115 Tankyrase
44	1521.2	39.9	4491	6	AR274526	AR274526 Sequence
45	1521.2	39.9	4491	6	AX467067	AX467067 Sequence

ALIGNMENTS

RESULT 1
AR392873
LOCUS AR392873 3816 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6617102.
ACCESSION AR392873
VERSION AR392873.1 GI:40118098
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3816)
AUTHORS Luo,Y., Chan,E., Xu,X., Huang,B. and Ossovskaya,V.
TITLE Tankyrase H, compositions involved in the cell cycle and methods of use
JOURNAL Patent: US 6617102-A 2 09-SEP-2003;

```
FEATURES             Location/Qualifiers
     source          1. .3816
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           100.0%;  Score 3816;  DB 6;  Length 3816;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3816;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560

Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680

Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740

Qy       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800

Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860

Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920

Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980

Qy       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
```

```
Qy       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100

Qy       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160

Qy       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220

Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280

Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340

Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400

Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460

Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520

Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
Db       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAA 3816
```

```
RESULT 2
AF309033
LOCUS       AF309033               6028 bp    mRNA    linear   PRI 31-JUL-2001
DEFINITION  Homo sapiens tankyrase-2 (TNKS-2) mRNA, complete cds.
ACCESSION   AF309033
VERSION     AF309033.1  GI:15042551
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6028)
  AUTHORS   Chi,N.-W. and Lodish,H.F.
  TITLE     A novel homolog of Tankyrase
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 6028)
  AUTHORS   Chi,N.-W. and Lodish,H.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-2000) Whitehead Institute for Biomedical
            Research, 9 Cambridge Center, Cambridge, MA 02142, USA
FEATURES             Location/Qualifiers
     source          1. .6028
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="10"
                     /map="10q23.1"
                     /note="STS WI-16054"
     gene            1. .6028
                     /gene="TNKS-2"
     CDS             252. .3752
                     /gene="TNKS-2"
                     /note="poly(ADP-ribose) polymerase containing ankyrin
                     repeats and SAM (sterile alpha module) motif"
                     /codon_start=1
                     /product="tankyrase-2"
                     /protein_id="AAK82330.1"
                     /db_xref="GI:15042552"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
     polyA_signal    6010. .6015
                     /gene="TNKS-2"
ORIGIN
```

Query Match 99.4%; Score 3793; DB 9; Length 6028;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3807; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 88

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 147

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 207

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 267

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 327

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 387

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 447

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 507

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 567

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 627

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 687

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 747

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 807

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 867

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 927

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 987

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1047

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1107

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1167

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1227

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1287

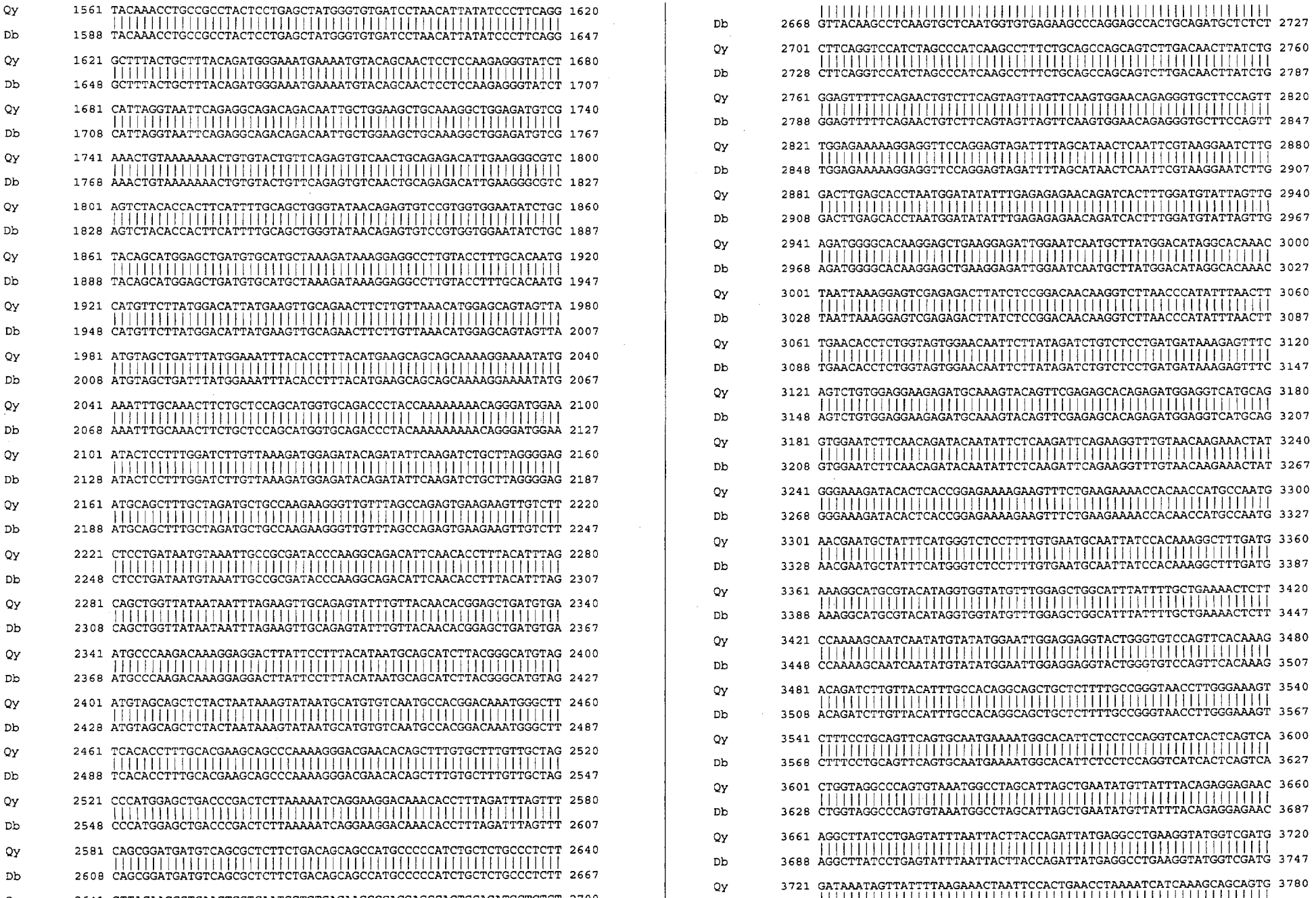
Qy       1261 GAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1347

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1407

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1467

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1527

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1587
```



```
Db       3748 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3807

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
              ||||||||||||||||||||||||||||||||||
Db       3808 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3840
```

RESULT 3
AX057579
LOCUS AX057579 6123 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 2 from Patent WO0077225.
ACCESSION AX057579
VERSION AX057579.1 GI:12310301
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chi,N.W. and Lodish,H.F.
TITLE A novel insulin signaling molecule
JOURNAL Patent: WO 0077225-A 2 21-DEC-2000;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; THE GENERAL
HOSPITAL CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .6123
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 99.4%; Score 3793; DB 6; Length 6123;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3807; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 88

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 147

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 207

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 267

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 327

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 387

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 447

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 507

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 567

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 627

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 687

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 747

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 807

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 867

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 927

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 987

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1047

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1107

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1167

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1227

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1287

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1347

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1407

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1467

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1527

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1587

Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1647

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1707
```

```
Qy      1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1708 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1767

Qy      1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1768 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1827

Qy      1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1828 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1887

Qy      1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1888 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1947

Qy      1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1948 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2007

Qy      1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2008 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2067

Qy      2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      2068 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2127

Qy      2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2128 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2187

Qy      2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2188 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2247

Qy      2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2248 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2307

Qy      2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2308 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2367

Qy      2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2368 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2427

Qy      2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2428 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2487

Qy      2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2488 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2547

Qy      2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2548 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2607

Qy      2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2608 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2667

Qy      2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2668 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2727

Qy      2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2728 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2787
```

```
Qy      2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2788 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2847

Qy      2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2848 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2907

Qy      2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2908 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2967

Qy      2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2968 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3027

Qy      3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3028 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3087

Qy      3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3088 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3147

Qy      3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3148 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3207

Qy      3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3208 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3267

Qy      3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3268 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3327

Qy      3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3328 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3387

Qy      3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3388 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3447

Qy      3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3448 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3507

Qy      3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3508 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3567

Qy      3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3568 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3627

Qy      3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3628 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3687

Qy      3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3688 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3747

Qy      3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3748 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3807

Qy      3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||
Db      3808 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3840
```

```
RESULT 4
AR361486
LOCUS       AR361486               4275 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 5 from patent US 6599728.
ACCESSION   AR361486
VERSION     AR361486.1  GI:33769334
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4275)
  AUTHORS   Morin,G.B., Funk,W.D. and Piatyszek,M.A.
  TITLE     Second mammalian tankyrase
  JOURNAL   Patent: US 6599728-A 5 29-JUL-2003
FEATURES             Location/Qualifiers
     source          1..4275
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           99.4%;  Score 3791.4;  DB 6;  Length 4275;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3806;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||  |  ||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAGGGGCGCGCG 120

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 179

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 239

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 299

Qy        241 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 359

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 419

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 479

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 539

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 599

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 659

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 719

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 779

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        780 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 839

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        840 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 899

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        900 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 959

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        960 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1019

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1020 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1079

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1080 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1139

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1199

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1259

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1319

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1379

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1439

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1440 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1499

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1500 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1559

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1560 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1619

Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1620 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1679

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1680 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1739

Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1740 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1799

Qy       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1800 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1859

Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1860 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1919
Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db       1920 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1979
Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db       1980 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2039
Qy       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db       2040 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2099
Qy       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db       2100 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2159
Qy       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db       2160 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2219
Qy       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db       2220 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2279
Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db       2280 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2339
Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db       2340 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2399
Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2400 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2459
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2460 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2519
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       2520 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2579
Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       2580 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2639
Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2640 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2699
Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2700 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2759
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2760 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2819
Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       2820 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2879
Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       2880 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2939
Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       2940 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2999
```

```
Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       3000 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3059
Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3060 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3119
Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3120 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3179
Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3180 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3239
Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3240 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3299
Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3300 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3359
Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3360 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3419
Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       3420 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3479
Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       3480 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3539
Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       3540 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3599
Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       3600 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3659
Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       3660 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3719
Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       3720 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3779
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       3780 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3839
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3840 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 5
AX062247
LOCUS AX062247 4406 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 106 from Patent WO0100849.
ACCESSION AX062247
VERSION AX062247.1 GI:12540148
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 106 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .4406
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .3790
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25062.1"
                     /db_xref="GI:12540149"
                     /db_xref="REMTREMBL:CAC25062"
                     /translation="ELAGGALPASAAASFQDPDGGFALPPPPRGSRGAGSPARGARGR
                     GHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRC
                     AGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHF
                     AAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDN
                     WNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE
                     LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVH
                     AKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLL
                     LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE
                     MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAH
                     NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL
                     QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQST
                     PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN
                     VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
                     GDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHG
                     ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQL
                     CALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSP
                     GATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDF
                     SITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLI
                     SGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY
                     NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAY
                     IGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFL
                     QFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG
                     "
ORIGIN

  Query Match             99.4%;  Score 3791.4;  DB 6;  Length 4406;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3806;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||  |  ||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 126

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 185

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 245

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 305

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 365

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 425

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 485

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 545

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 605

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 665

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 725

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 785

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 845

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 905

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 965

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1025

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1085

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1145

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1205

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1265

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1325

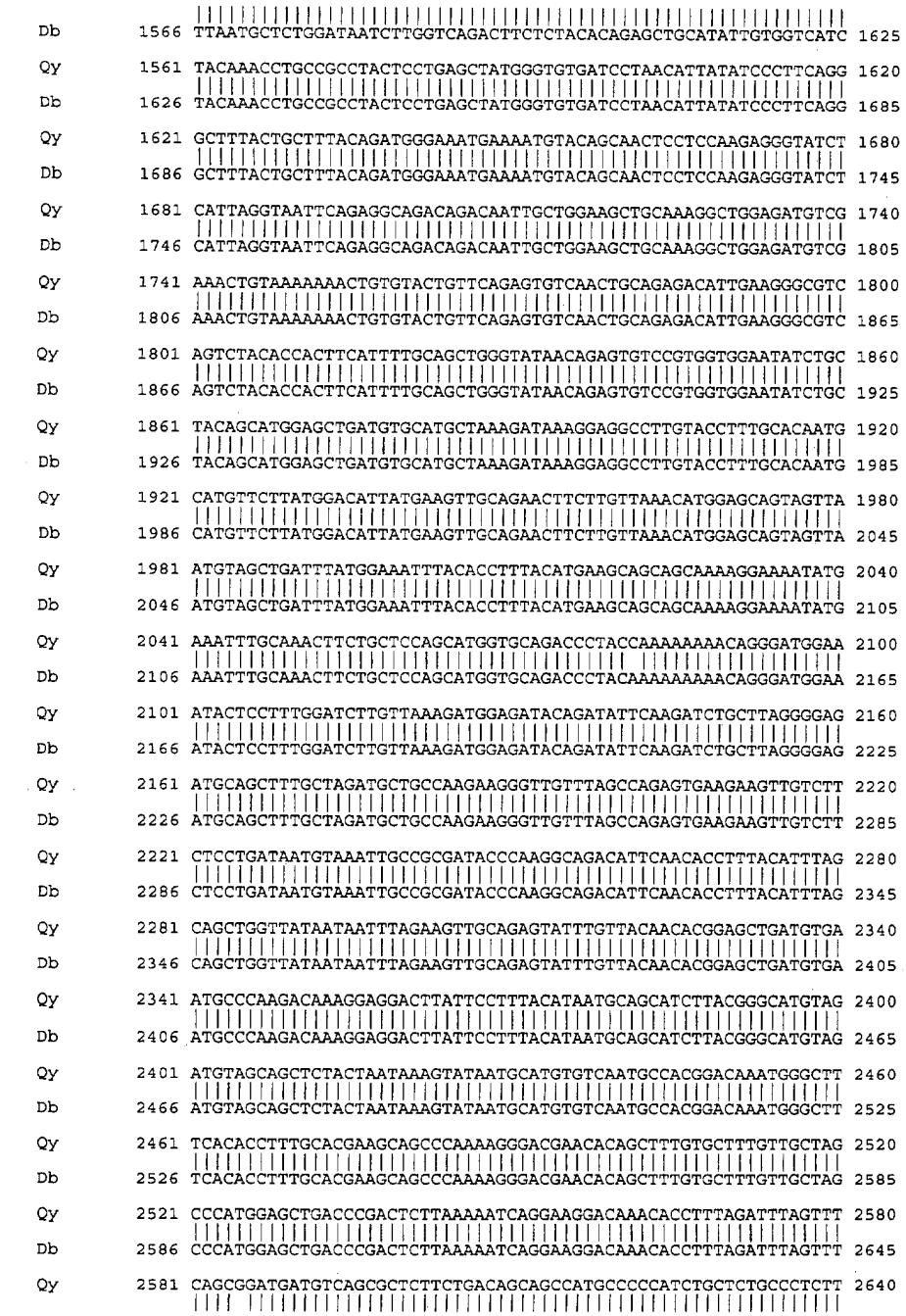
Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1385

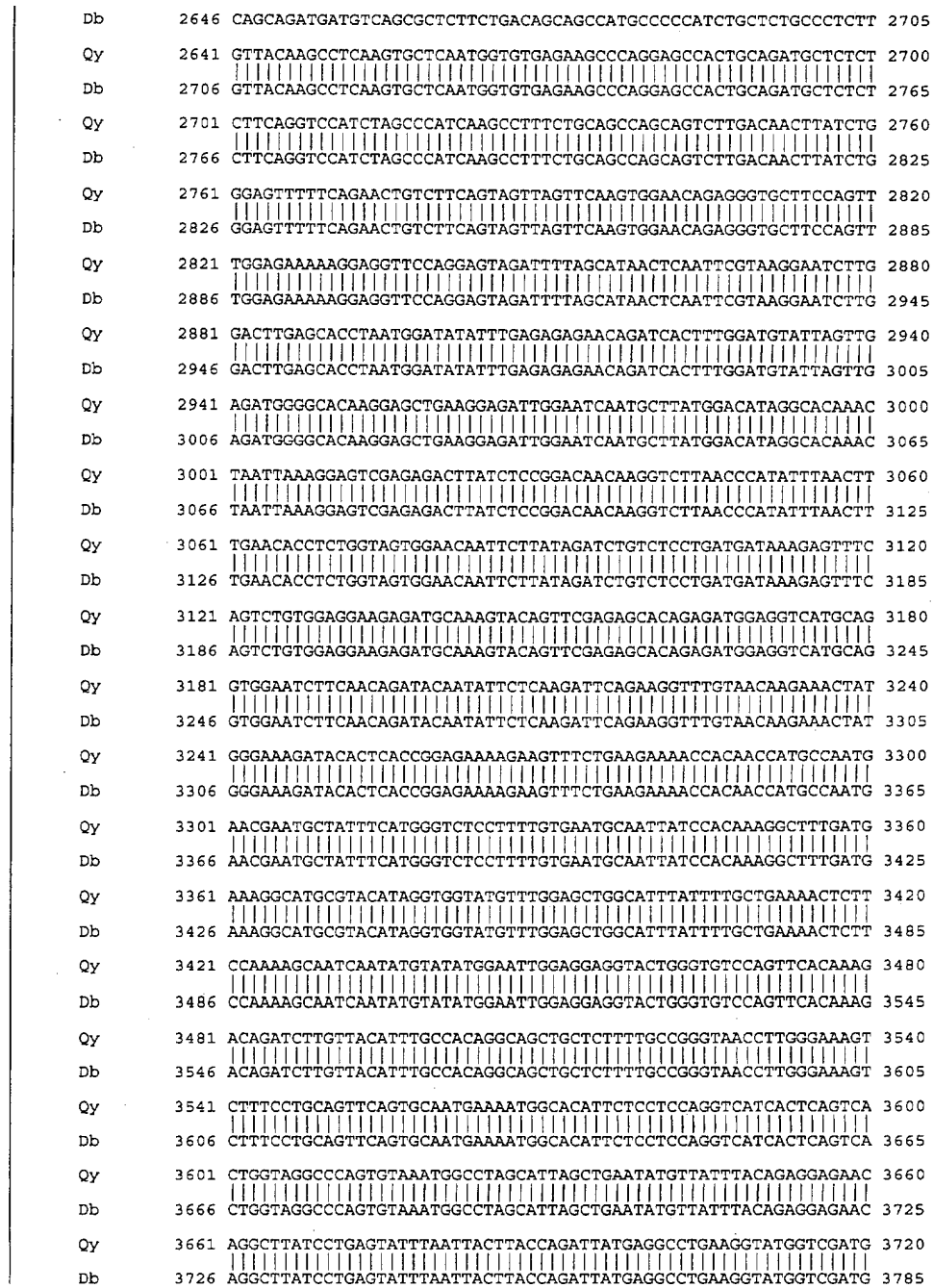
Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1445

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1505

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1565

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
```


Db 1566 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1625
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 1626 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1685
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 1686 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1745
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 1746 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1805
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1806 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1865
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 1866 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1925
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 1926 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1985
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 1986 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2045
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2046 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2105
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db 2106 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2165
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2166 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2225
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2226 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2285
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2286 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2345
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db 2346 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2405
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 2406 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2465
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db 2466 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2525
Qy 2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db 2526 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2585
Qy 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 2586 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2645
Qy 2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640


Db 2646 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2705
Qy 2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db 2706 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2765
Qy 2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db 2766 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2825
Qy 2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db 2826 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2885
Qy 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 2886 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2945
Qy 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db 2946 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3005
Qy 2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db 3006 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3065
Qy 3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db 3066 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3125
Qy 3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db 3126 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3185
Qy 3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db 3186 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3245
Qy 3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db 3246 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3305
Qy 3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db 3306 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3365
Qy 3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db 3366 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3425
Qy 3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db 3426 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3485
Qy 3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db 3486 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3545
Qy 3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db 3546 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3605
Qy 3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db 3606 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3665
Qy 3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db 3666 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3725
Qy 3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db 3726 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3785

```
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3786 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3845

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             |||||||||||||||||||||||||||||||||
Db       3846 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3878


RESULT 6
AX062275
LOCUS       AX062275                4992 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 134 from Patent WO0100849.
ACCESSION   AX062275
VERSION     AX062275.1  GI:12540176
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 134 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .4992
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             876. .4376
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25067.1"
                     /db_xref="GI:12540177"
                     /db_xref="REMTREMBL:CAC25067"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match           99.4%;  Score 3791.4;  DB 6;  Length 4992;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 3806;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 712

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 771

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 831

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 891

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 951

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 1011

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1012 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 1071

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1072 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 1131

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1132 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 1191

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1192 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 1251

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 1311

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 1371

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1372 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 1431

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1432 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 1491

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1492 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1551

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1611

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1671

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1672 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1731

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1732 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1791

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1792 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1851

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1852 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1911
Qy      1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db      1912 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1971
Qy      1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db      1972 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 2031
Qy      1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db      2032 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 2091
Qy      1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db      2092 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 2151
Qy      1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db      2152 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 2211
Qy      1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db      2212 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 2271
Qy      1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db      2272 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 2331
Qy      1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db      2332 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 2391
Qy      1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db      2392 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 2451
Qy      1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db      2452 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2511
Qy      1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db      2512 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2571
Qy      1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db      2572 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2631
Qy      1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db      2632 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2691
Qy      2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db      2692 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2751
Qy      2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db      2752 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2811
Qy      2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db      2812 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2871
Qy      2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db      2872 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2931
Qy      2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db      2932 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2991

Qy      2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db      2992 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3051
Qy      2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db      3052 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 3111
Qy      2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db      3112 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 3171
Qy      2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db      3172 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3231
Qy      2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db      3232 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3291
Qy      2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db      3292 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 3351
Qy      2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db      3352 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 3411
Qy      2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db      3412 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 3471
Qy      2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db      3472 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3531
Qy      2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db      3532 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3591
Qy      2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db      3592 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3651
Qy      3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db      3652 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3711
Qy      3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db      3712 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3771
Qy      3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db      3772 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3831
Qy      3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db      3832 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3891
Qy      3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db      3892 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3951
Qy      3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db      3952 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 4011
Qy      3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db      4012 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 4071
```

```
Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4072 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 4131

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4132 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 4191

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4192 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 4251

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4252 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 4311

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4312 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 4371

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4372 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 4431

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||
Db       4432 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4464
```

```
RESULT 7
AX062273
LOCUS       AX062273                5002 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 132 from Patent WO0100849.
ACCESSION   AX062273
VERSION     AX062273.1  GI:12540174
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 132 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .5002
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             229. .4386
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25066.1"
                     /db_xref="GI:12540175"
                     /db_xref="REMTREMBL:CAC25066"
                     /translation="MWKLKTRTISTVRLPKECARSEQNFQGLFFSVMVVRKFSLESPG
                     LKTTTTKNTICRIVRLQQNPPQRWRWDEAPSPAAEASRLTFPTNPSRRLASRNLPSRC
                     PATAHARDDVTCAPGAGRSWQELAGGALPASAAASFQDPDGGFALPPPPRGSRGAGSP
                     ARGARGRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAAR
                     IMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR
                     KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGA
                     DPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLT
                     GEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL
                     QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR
                     VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRI
                     KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLH
                     VASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIS
                     LQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD
                     IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLV
                     KHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT
                     DIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVA
                     EYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA
                     QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQV
                     LNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKK
                     EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLI
                     KGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHA
                     GGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKG
                     FDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRV
                     TLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMR
                     PEGMVDG"
ORIGIN

  Query Match           99.4%;  Score 3791.4;  DB 6;  Length 5002;
  Best Local Similarity 99.8%;  Pred. No. 0;
  Matches 3806;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 722

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 781

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 841

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 901

Qy        241 GCGCCGGCGGGGGAGCCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 961

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 1021

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 1081

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 1141

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 1201

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 1261

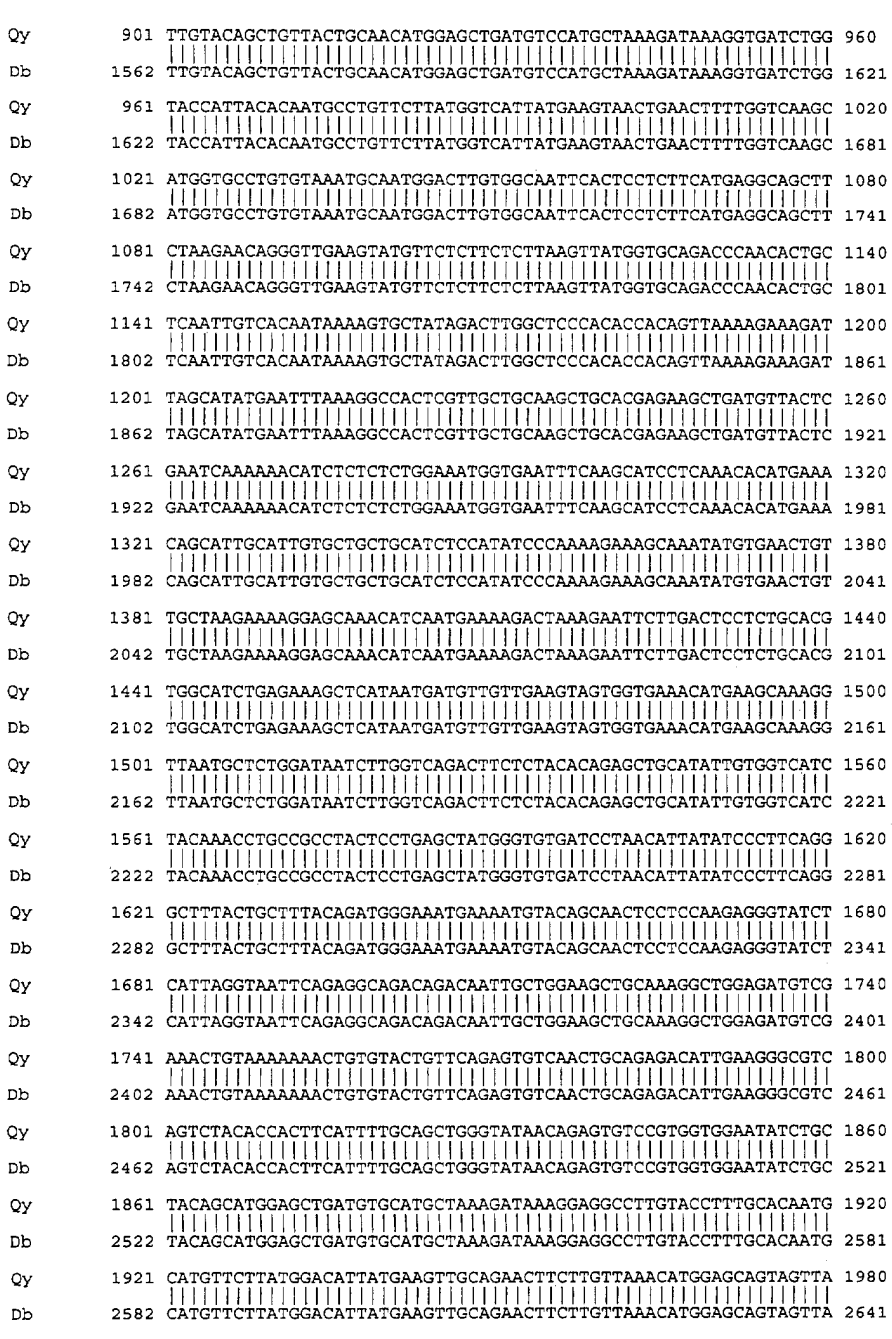
Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 1321

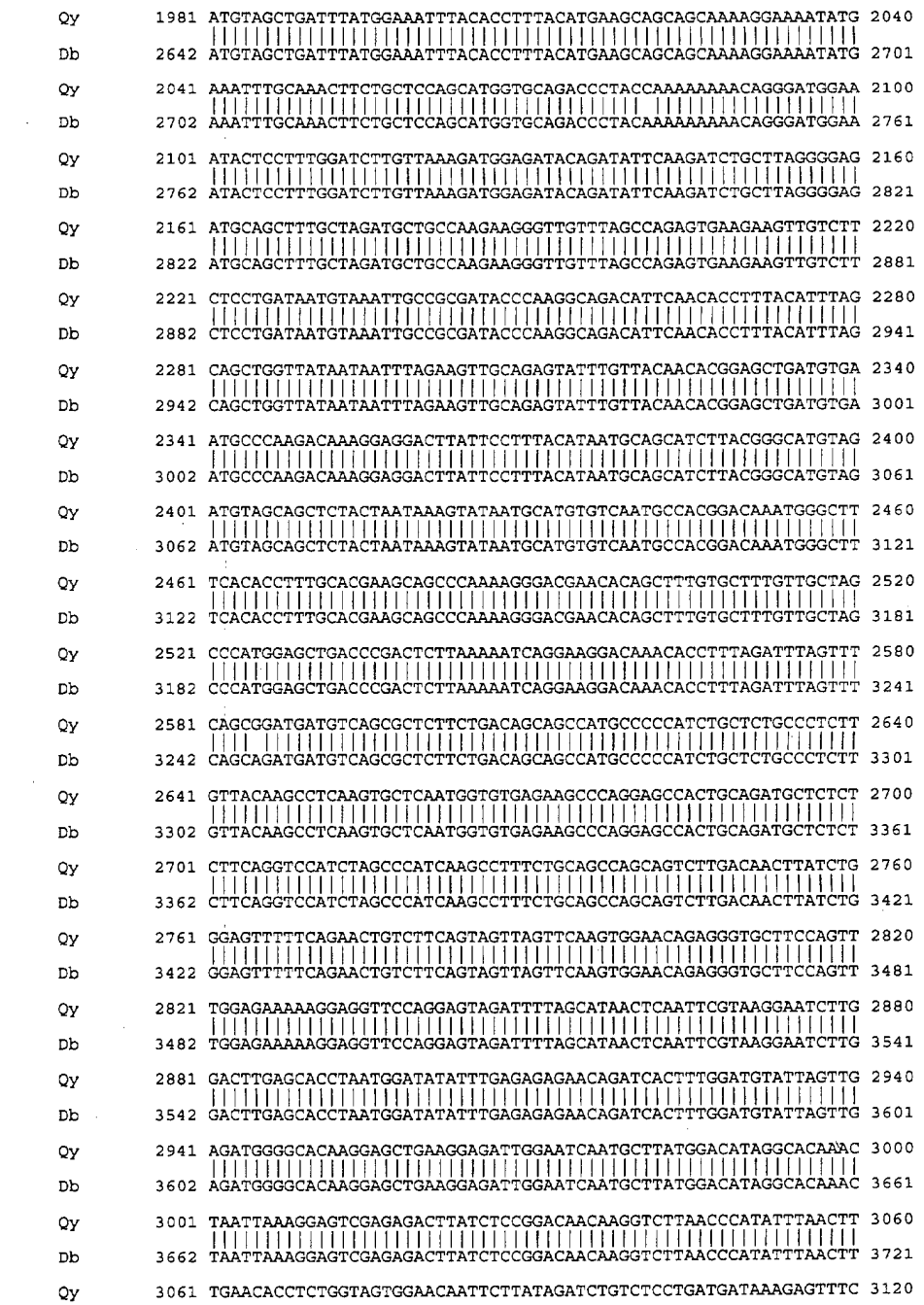
Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 1381

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 1441

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 1501

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1561
```


Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 1562 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1621
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1622 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1681
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1682 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1741
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1742 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1801
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1802 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1861
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1862 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1921
Qy 1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1922 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1981
Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1982 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 2041
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 2042 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 2101
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 2102 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 2161
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 2162 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 2221
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 2222 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 2281
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 2282 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 2341
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 2342 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 2401
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 2402 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 2461
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 2462 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2521
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 2522 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2581
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 2582 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2641


Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2642 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2701
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db 2702 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2761
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2762 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2821
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2822 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2881
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2882 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2941
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db 2942 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 3001
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 3002 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3061
Qy 2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db 3062 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 3121
Qy 2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db 3122 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 3181
Qy 2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db 3182 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3241
Qy 2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db 3242 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3301
Qy 2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db 3302 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 3361
Qy 2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db 3362 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 3421
Qy 2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db 3422 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 3481
Qy 2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db 3482 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3541
Qy 2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db 3542 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3601
Qy 2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db 3602 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3661
Qy 3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db 3662 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3721
Qy 3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120

```
Db      3722 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3781

Qy      3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180

Db      3782 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3841

Qy      3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240

Db      3842 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3901

Qy      3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300

Db      3902 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3961

Qy      3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360

Db      3962 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 4021

Qy      3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420

Db      4022 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 4081

Qy      3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480

Db      4082 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 4141

Qy      3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540

Db      4142 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 4201

Qy      3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600

Db      4202 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 4261

Qy      3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660

Db      4262 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 4321

Qy      3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720

Db      4322 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 4381

Qy      3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780

Db      4382 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 4441

Qy      3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813

Db      4442 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4474
```

```
RESULT 8
AF438201
LOCUS       AF438201               6189 bp    mRNA    linear   PRI 12-DEC-2001
DEFINITION  Homo sapiens tankyrase II mRNA, complete cds.
ACCESSION   AF438201
VERSION     AF438201.1  GI:17530294
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6189)
  AUTHORS   Yin,Y. and Gelmann,E.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) Lombardi Cancer Center, Georgetown
            University, 3800 Reservoir Rd, NW, Washington, DC 20007, USA
FEATURES             Location/Qualifiers
     source          1. .6189
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q24"
     CDS             380. .3880
                     /note="telomerase binding protein"
                     /codon_start=1
                     /product="tankyrase II"
                     /protein_id="AAL40795.1"
                     /db_xref="GI:17530295"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match           99.4%;  Score 3791.4;  DB 9;  Length 6189;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 3806;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Db        157 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 216

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Db        217 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 275

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Db        276 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 335

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240

Db        336 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 395

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300

Db        396 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 455

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360

Db        456 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 515

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420

Db        516 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 575

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480

Db        576 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 635

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540

Db        636 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 695

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
```

```
Db        696 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 755
Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db        756 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 815
Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db        816 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 875
Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db        876 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 935
Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db        936 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 995
Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db        996 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1055
Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db       1056 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1115
Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db       1116 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1175
Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db       1176 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1235
Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db       1236 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1295
Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db       1296 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1355
Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db       1356 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1415
Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db       1416 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1475
Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db       1476 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1535
Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db       1536 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1595
Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db       1596 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1655
Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db       1656 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1715
Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db       1716 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1775
Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db       1776 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1835

Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db       1836 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1895
Qy       1741 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db       1896 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1955
Qy       1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db       1956 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2015
Qy       1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db       2016 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2075
Qy       1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db       2076 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2135
Qy       1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db       2136 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2195
Qy       2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2100
Db       2196 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2255
Qy       2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db       2256 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2315
Qy       2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db       2316 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2375
Qy       2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db       2376 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2435
Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db       2436 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2495
Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2496 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2555
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2556 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2615
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       2616 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2675
Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       2676 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2735
Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2736 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2795
Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2796 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2855
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2856 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2915
```

```
Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2916 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2975

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2976 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3035

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3036 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3095

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3096 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3155

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3156 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3215

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3216 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3275

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3276 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3335

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3336 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3395

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3396 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3455

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3456 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3515

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3516 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3575

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3576 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3635

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3636 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3695

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3696 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3755

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3756 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3815

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3816 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3875

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3876 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3935

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||
Db       3936 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3968
```

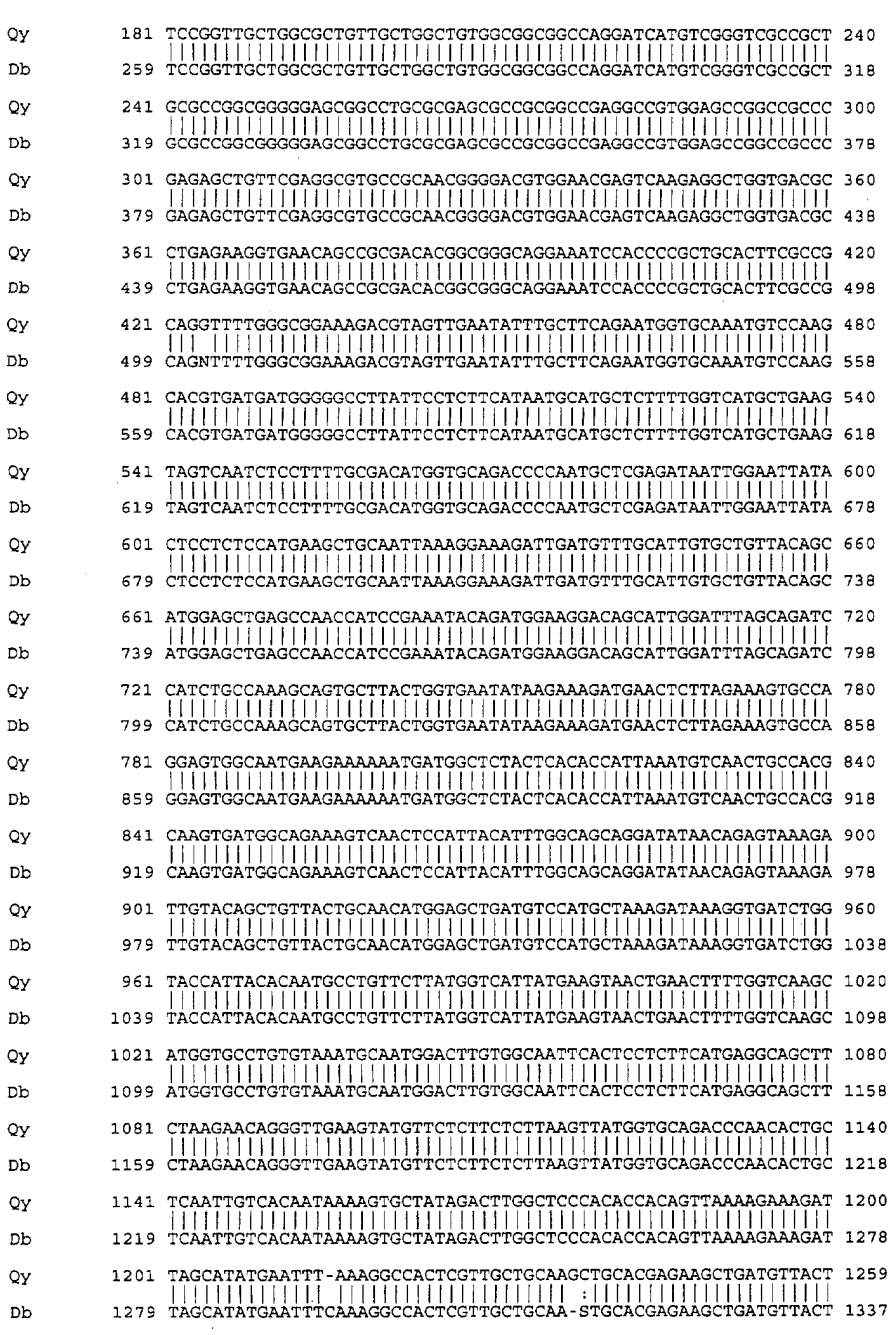
```
RESULT 9
AF305081
LOCUS       AF305081                4295 bp    mRNA    linear   PRI 17-APR-2001
DEFINITION  Homo sapiens tankyrase-related protein mRNA, partial cds.
ACCESSION   AF305081
VERSION     AF305081.1  GI:10953951
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4295)
  AUTHORS   Monz,D., Munnia,A., Comtesse,N., Fischer,U., Steudel,W.I.,
            Feiden,W., Glass,B. and Meese,E.U.
  TITLE     Novel tankyrase-related gene detected with meningioma-specific sera
  JOURNAL   Clin. Cancer Res. 7 (1), 113-119 (2001)
  MEDLINE   21072550
   PUBMED   11205898
REFERENCE   2  (bases 1 to 4295)
  AUTHORS   Monz,D.W. and Meese,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Human Genetics, Universitaet des
            Saarlandes, Geb. 60, Homburg 66421, Germany
FEATURES             Location/Qualifiers
     source          1. .4295
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             <1. .3800
                     /codon_start=3
                     /product="tankyrase-related protein"
                     /protein_id="AAG25674.1"
                     /db_xref="GI:10953952"
                     /translation="HEAGRSWQEGPCQLPPPRRFRTRTADSRCLRRRGAAGGQGAQRG
                     ARVGGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMS
                     GRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST
                     PLHFAAXFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPN
                     ARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEY
                     KKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHG
                     ADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV
                     CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFQRPLVAAXAREADVTRIKKH
                     LSLEMVIQASQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASE
                     KAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGF
                     TALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR
                     QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA
                     VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD
                     LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLL
                     QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGR
                     TQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGV
                     RSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPG
                     VDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVE
                     RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIF
                     NRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDER
                     HAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGK
                     SFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGM
                     VDG"
ORIGIN

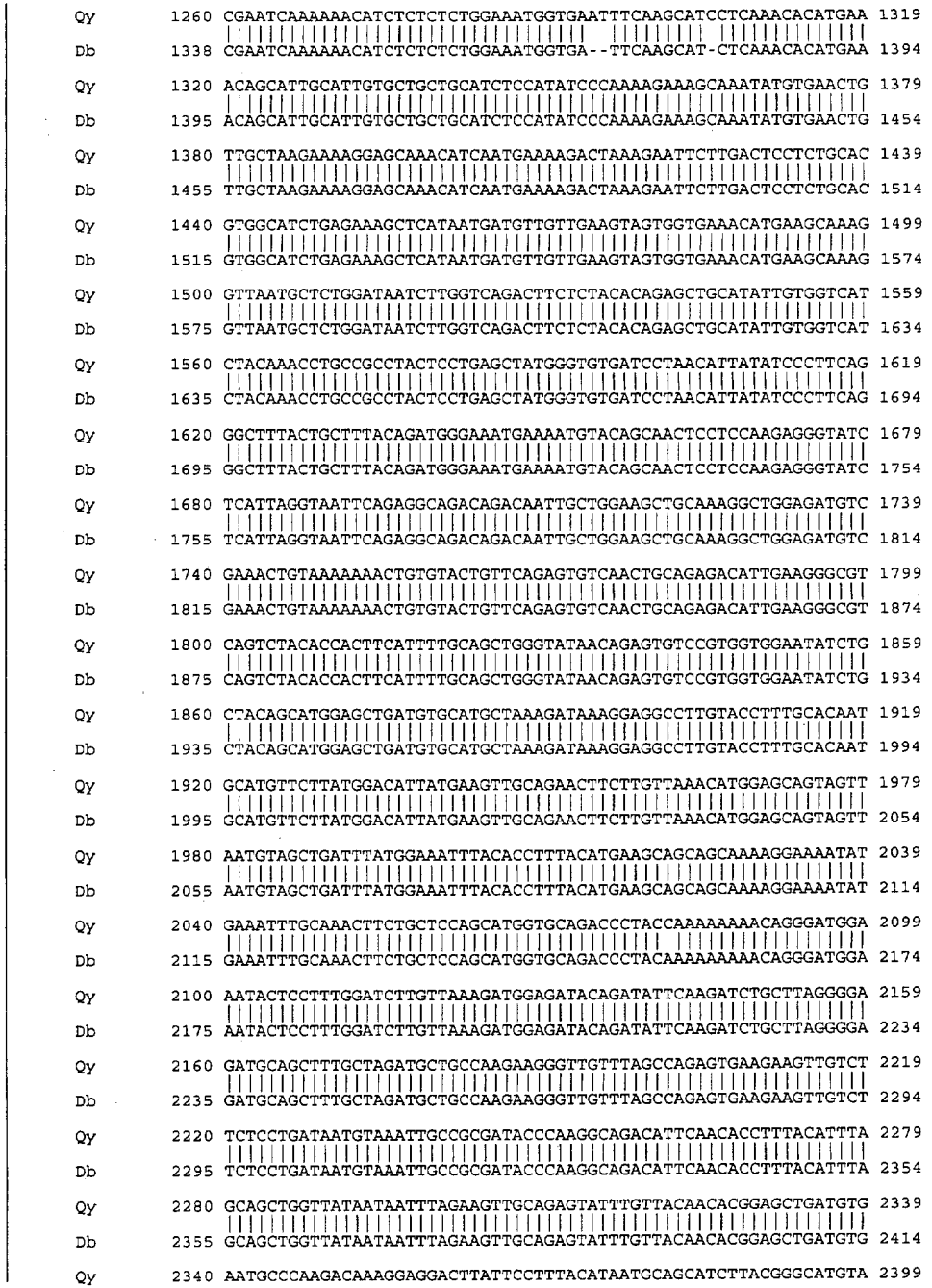
  Query Match             97.9%;  Score 3737;  DB 9;  Length 4295;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 3798;  Conservative    1;  Mismatches    7;  Indels    8;  Gaps    5;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||  |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 141

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              |||||  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 TGGGC---GGCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 198

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 258
```


Qy 181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
Db 259 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 318
Qy 241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db 319 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 378
Qy 301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db 379 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 438
Qy 361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db 439 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 498
Qy 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db 499 CAGNTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 558
Qy 481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db 559 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 618
Qy 541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db 619 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 678
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db 679 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 738
Qy 661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db 739 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 798
Qy 721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db 799 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 858
Qy 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db 859 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 918
Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 919 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 978
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 979 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1038
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1039 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1098
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1099 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1158
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1159 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1218
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1219 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1278
Qy 1201 TAGCATATGAATTT-AAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACT 1259
Db 1279 TAGCATATGAATTTCAAAGGCCACTCGTTGCTGCAA-STGCACGAGAAGCTGATGTTACT 1337


Qy 1260 CGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAA 1319
Db 1338 CGAATCAAAAAACATCTCTCTCTGGAAATGGTGA--TTCAAGCAT-CTCAAACACATGAA 1394
Qy 1320 ACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTG 1379
Db 1395 ACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTG 1454
Qy 1380 TTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCAC 1439
Db 1455 TTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCAC 1514
Qy 1440 GTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAG 1499
Db 1515 GTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAG 1574
Qy 1500 GTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCAT 1559
Db 1575 GTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCAT 1634
Qy 1560 CTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAG 1619
Db 1635 CTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAG 1694
Qy 1620 GGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATC 1679
Db 1695 GGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATC 1754
Qy 1680 TCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTC 1739
Db 1755 TCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTC 1814
Qy 1740 GAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGT 1799
Db 1815 GAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGT 1874
Qy 1800 CAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTG 1859
Db 1875 CAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTG 1934
Qy 1860 CTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAAT 1919
Db 1935 CTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAAT 1994
Qy 1920 GCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTT 1979
Db 1995 GCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTT 2054
Qy 1980 AATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATAT 2039
Db 2055 AATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATAT 2114
Qy 2040 GAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGA 2099
Db 2115 GAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGA 2174
Qy 2100 AATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGA 2159
Db 2175 AATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGA 2234
Qy 2160 GATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCT 2219
Db 2235 GATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCT 2294
Qy 2220 TCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTA 2279
Db 2295 TCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTA 2354
Qy 2280 GCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTG 2339
Db 2355 GCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTG 2414
Qy 2340 AATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTA 2399

```
Db      2415 AATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTA 2474
Qy      2400 GATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCT 2459
Db      2475 GATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCT 2534
Qy      2460 TTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTA 2519
Db      2535 TTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTA 2594
Qy      2520 GCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTT 2579
Db      2595 GCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTT 2654
Qy      2580 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2639
Db      2655 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2714
Qy      2640 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2699
Db      2715 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2774
Qy      2700 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2759
Db      2775 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2834
Qy      2760 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2819
Db      2835 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2894
Qy      2820 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2879
Db      2895 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2954
Qy      2880 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2939
Db      2955 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 3014
Qy      2940 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 2999
Db      3015 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 3074
Qy      3000 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3059
Db      3075 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3134
Qy      3060 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3119
Db      3135 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3194
Qy      3120 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3179
Db      3195 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3254
Qy      3180 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3239
Db      3255 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3314
Qy      3240 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3299
Db      3315 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3374
Qy      3300 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3359
Db      3375 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3434
Qy      3360 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3419
Db      3435 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3494
Qy      3420 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3479
Db      3495 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3554
Qy      3480 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3539
Db      3555 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3614
Qy      3540 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3599
Db      3615 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3674
Qy      3600 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAA 3659
Db      3675 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAA 3734
Qy      3660 CAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGAT 3719
Db      3735 CAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGAT 3794
Qy      3720 GGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGT 3779
Db      3795 GGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGT 3854
Qy      3780 GGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db      3855 GGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3888
```

```
RESULT 10
AF329696
LOCUS       AF329696               3815 bp    mRNA    linear   PRI 14-MAY-2001
DEFINITION  Homo sapiens tankyrase 2 mRNA, complete cds.
ACCESSION   AF329696
VERSION     AF329696.1  GI:13161041
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3815)
  AUTHORS   Lyons,R.J., Deane,R., Lynch,D.K., Ye,Z.S., Sanderson,G.M.,
            Eyre,H.J., Sutherland,G.R. and Daly,R.J.
  TITLE     Identification of a novel human tankyrase through its interaction
            with the adaptor protein Grb14
  JOURNAL   J. Biol. Chem. 276 (20), 17172-17180 (2001)
  MEDLINE   21264473
   PUBMED   11278563
REFERENCE   2  (bases 1 to 3815)
  AUTHORS   Lyons,R.J., Deane,R., Lynch,D.K., Ye,Z.-S.J., Sanderson,G.M.,
            Eyre,H.J., Sutherland,G.R. and Daly,R.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2000) Cancer Research, Garvan Institute of
            Medical Research, 384 Victoria St., Sydney, NSW 2010, Australia
FEATURES             Location/Qualifiers
     source          1. .3815
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q23.2"
     CDS             140. .3640
                     /codon_start=1
                     /product="tankyrase 2"
                     /protein_id="AAK13463.1"
                     /db_xref="GI:13161042"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
```

KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
AYPEYLITYQIMRPEGMVDG"

ORIGIN

Query Match 97.6%; Score 3724.8; DB 9; Length 3815;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3726; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        86 GGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCCTGAGCGCGTCTTCTCCGGGGGG 145
Db         1 GGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCCTGAGCGCGTCTTCTCCGGGGGG 60

Qy       146 CCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCCGGTTGCTGGCGCTGTTGCTGGC 205
Db        61 CCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCCGGTTGCTGGCGCTGTTGCTGGC 120

Qy       206 TGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGC 265
Db       121 TGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGC 180

Qy       266 GAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAA 325
Db       181 GAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAA 240

Qy       326 CGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACAC 385
Db       241 CGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACAC 300

Qy       386 GGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGT 445
Db       301 GGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGT 360

Qy       446 TGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCC 505
Db       361 TGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCC 420

Qy       506 TCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGG 565
Db       421 TCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGG 480

Qy       566 TGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAA 625
Db       481 TGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAA 540

Qy       626 AGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAA 685
Db       541 AGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAA 600

Qy       686 TACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGG 745
Db       601 TACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGG 660

Qy       746 TGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGAT 805
Db       661 TGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGAT 720

Qy       806 GGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCC 865
Db       721 GGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCC 780

Qy       866 ATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGG 925
Db       781 ATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGG 840

Qy       926 AGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTA 985
Db       841 AGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTA 900

Qy       986 TGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGA 1045
Db       901 TGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGA 960

Qy      1046 CTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTC 1105
Db       961 CTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTC 1020

Qy      1106 TCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTAT 1165
Db       1021 TCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTAT 1080

Qy      1166 AGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTC 1225
Db       1081 AGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTC 1140

Qy      1226 GTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGA 1285
Db       1141 GTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGA 1200

Qy      1286 AATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATC 1345
Db       1201 AATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATC 1260

Qy      1346 TCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAA 1405
Db       1261 TCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAA 1320

Qy      1406 TGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGA 1465
Db       1321 TGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGA 1380

Qy      1466 TGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCA 1525
Db       1381 TGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCA 1440

Qy      1526 GACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAG 1585
Db       1441 GACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAG 1500

Qy      1586 CTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAA 1645
Db       1501 CTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAA 1560

Qy      1646 TGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAG 1705
Db       1561 TGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAG 1620

Qy      1706 ACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGT 1765
Db       1621 ACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGT 1680

Qy      1766 TCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGC 1825
Db       1681 TCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGC 1740

Qy      1826 TGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGC 1885
Db       1741 TGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGC 1800

Qy      1886 TAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGT 1945
Db       1801 TAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGT 1860

Qy      1946 TGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTAC 2005
Db       1861 TGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTAC 1920
```

```
Qy       2006 ACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCA 2065
Db       1921 ACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCA 1980

Qy       2066 TGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGA 2125
Db       1981 TGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGA 2040

Qy       2126 TGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAA 2185
Db       2041 TGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAA 2100

Qy       2186 GAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGA 2245
Db       2101 GAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGA 2160

Qy       2246 TACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGT 2305
Db       2161 TACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGT 2220

Qy       2306 TGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTAT 2365
Db       2221 TGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTAT 2280

Qy       2366 TCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTA 2425
Db       2281 TCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTA 2340

Qy       2426 TAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCA 2485
Db       2341 TAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCA 2400

Qy       2486 AAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAA 2545
Db       2401 AAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAA 2460

Qy       2546 AAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCT 2605
Db       2461 AAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCT 2520

Qy       2606 GACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGG 2665
Db       2521 GACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGG 2580

Qy       2666 TGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAG 2725
Db       2581 TGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAG 2640

Qy       2726 CCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGT 2785
Db       2641 CCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGT 2700

Qy       2786 AGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGT 2845
Db       2701 AGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGT 2760

Qy       2846 AGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATT 2905
Db       2761 AGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATT 2820

Qy       2906 TGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGA 2965
Db       2821 TGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGA 2880

Qy       2966 GATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTAT 3025
Db       2881 GATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTAT 2940

Qy       3026 CTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAAT 3085
Db       2941 CTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAAT 3000

Qy       3086 TCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAG 3145
Db       3001 TCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAG 3060

Qy       3146 TACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATAT 3205
Db       3061 TACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATAT 3120

Qy       3206 TCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAA 3265
Db       3121 TCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAA 3180

Qy       3266 AGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCC 3325
Db       3181 AGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCC 3240

Qy       3326 TTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTAT 3385
Db       3241 TTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTAT 3300

Qy       3386 GTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGG 3445
Db       3301 GTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGG 3360

Qy       3446 AATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAG 3505
Db       3361 AATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAG 3420

Qy       3506 GCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAA 3565
Db       3421 GCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAA 3480

Qy       3566 AATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCT 3625
Db       3481 AATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCT 3540

Qy       3626 AGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAC 3685
Db       3541 AGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAC 3600

Qy       3686 TTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTA 3745
Db       3601 TTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTA 3660

Qy       3746 ATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3805
Db       3661 ATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3720

Qy       3806 AAAAAAAA 3813
Db       3721 AAAAAAAA 3728
```

```
RESULT 11
AX062241
LOCUS       AX062241               4127 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 100 from Patent WO0100849.
ACCESSION   AX062241
VERSION     AX062241.1  GI:12540142
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 100 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .4127
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

```
     CDS             <2. .3511
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25060.1"
                     /db_xref="GI:12540143"
                     /db_xref="REMTREMBL:CAC25060"
                     /translation="ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVK
                     RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNAC
                     SFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG
                     RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPL
                     HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM
                     DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFK
                     GHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK
                     GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ
                     TCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV
                     ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPL
                     HNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKK
                     NRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRH
                     STPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNAC
                     VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLT
                     AAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSS
                     VVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKE
                     LKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE
                     EEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER
                     MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK
                     DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYR
                     GEQAYPEYLITYQIMRPEGMVDG"
     3'UTR           3509. .4127
ORIGIN

  Query Match          94.2%;  Score 3595.8;  DB 6;  Length 4127;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3597;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600

Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840

Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900

Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960

Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020

Qy       1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080

Qy       1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140

Qy       1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200

Qy       1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260

Qy       1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320

Qy       1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380

Qy       1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440

Qy       1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500

Qy       1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560

Qy       1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1620

Qy       1835 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680
```

```
Qy       1895 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1954
Db       1681 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1740

Qy       1955 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 2014
Db       1741 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 1800

Qy       2015 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
Db       1801 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860

Qy       2075 CCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 2134
Db       1861 CCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 1920

Qy       2135 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
Db       1921 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980

Qy       2195 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2254
Db       1981 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2040

Qy       2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2314
Db       2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2100

Qy       2315 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
Db       2101 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160

Qy       2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2434
Db       2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2220

Qy       2435 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
Db       2221 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280

Qy       2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
Db       2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340

Qy       2555 AGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
Db       2341 AGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGC 2400

Qy       2615 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2674
Db       2401 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2460

Qy       2675 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2734
Db       2461 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2520

Qy       2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
Db       2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580

Qy       2795 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2854
Db       2581 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2640

Qy       2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2914
Db       2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2700

Qy       2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2974
Db       2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2760

Qy       2975 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 3034
Db       2761 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 2820

Qy       3035 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 3094
Db       2821 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 2880

Qy       3095 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 3154
Db       2881 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 2940

Qy       3155 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3214
Db       2941 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3000

Qy       3215 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3274
Db       3001 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3060

Qy       3275 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3334
Db       3061 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3120

Qy       3335 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3394
Db       3121 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3180

Qy       3395 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3454
Db       3181 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3240

Qy       3455 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3514
Db       3241 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3300

Qy       3515 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3574
Db       3301 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3360

Qy       3575 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3634
Db       3361 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3420

Qy       3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3694
Db       3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3480

Qy       3695 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3754
Db       3481 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3540

Qy       3755 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3541 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3599
```

RESULT 12
AR232161
LOCUS AR232161 4512 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6455290.
ACCESSION AR232161
VERSION AR232161.1 GI:27274048
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4512)
AUTHORS Berthelsen,J., Toma,S. and Isacchi,A.
TITLE Tankyrase homolog protein (THP), nucleic acids, and methods related to the same
JOURNAL Patent: US 6455290-A 3 24-SEP-2002;
FEATURES Location/Qualifiers
source 1. .4512

```
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match          94.2%;  Score 3594.8;  DB 6;  Length 4512;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3593;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       210 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         8 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 67

Qy       270 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 329
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        68 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 127

Qy       330 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 389
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       128 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 187

Qy       390 GGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 449
            |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       188 GGCAGGAAATCCACCCCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 247

Qy       450 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 509
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       248 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 307

Qy       510 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 569
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       308 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 367

Qy       570 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 629
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       368 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 427

Qy       630 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 689
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       428 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 487

Qy       690 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 749
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       488 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 547

Qy       750 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 809
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       548 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 607

Qy       810 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 869
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       608 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 667

Qy       870 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 929
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       668 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 727

Qy       930 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 989
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       728 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 787

Qy       990 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 1049
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       788 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 847

Qy      1050 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 1109
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       848 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 907

Qy      1110 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 1169
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       908 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 967

Qy      1170 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       968 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027

Qy      1230 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCTCTGGAAATG 1289
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1028 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1087

Qy      1290 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCA 1349
            |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       1088 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCA 1147

Qy      1350 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAA 1409
            ||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1148 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAA 1207

Qy      1410 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTT 1469
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||
Db       1208 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTT 1267

Qy      1470 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1529
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1327

Qy      1530 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1589
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1328 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1387

Qy      1590 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1649
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1447

Qy      1650 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1709
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1507

Qy      1710 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1769
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1508 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1567

Qy      1770 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1829
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1568 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1627

Qy      1830 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1889
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1628 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1687

Qy      1890 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1949
            |||||||::|:|||||||||||||||||||||||||||||||||||||||||||||||||
Db       1688 GATAAAGRRGSCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1747

Qy      1950 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 2009
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1748 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 1807

Qy      2010 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1808 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 1867

Qy      2070 GCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 2129
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1868 GCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 1927

Qy      2130 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 2189
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1928 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 1987

Qy      2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1988 AGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2047

Qy      2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

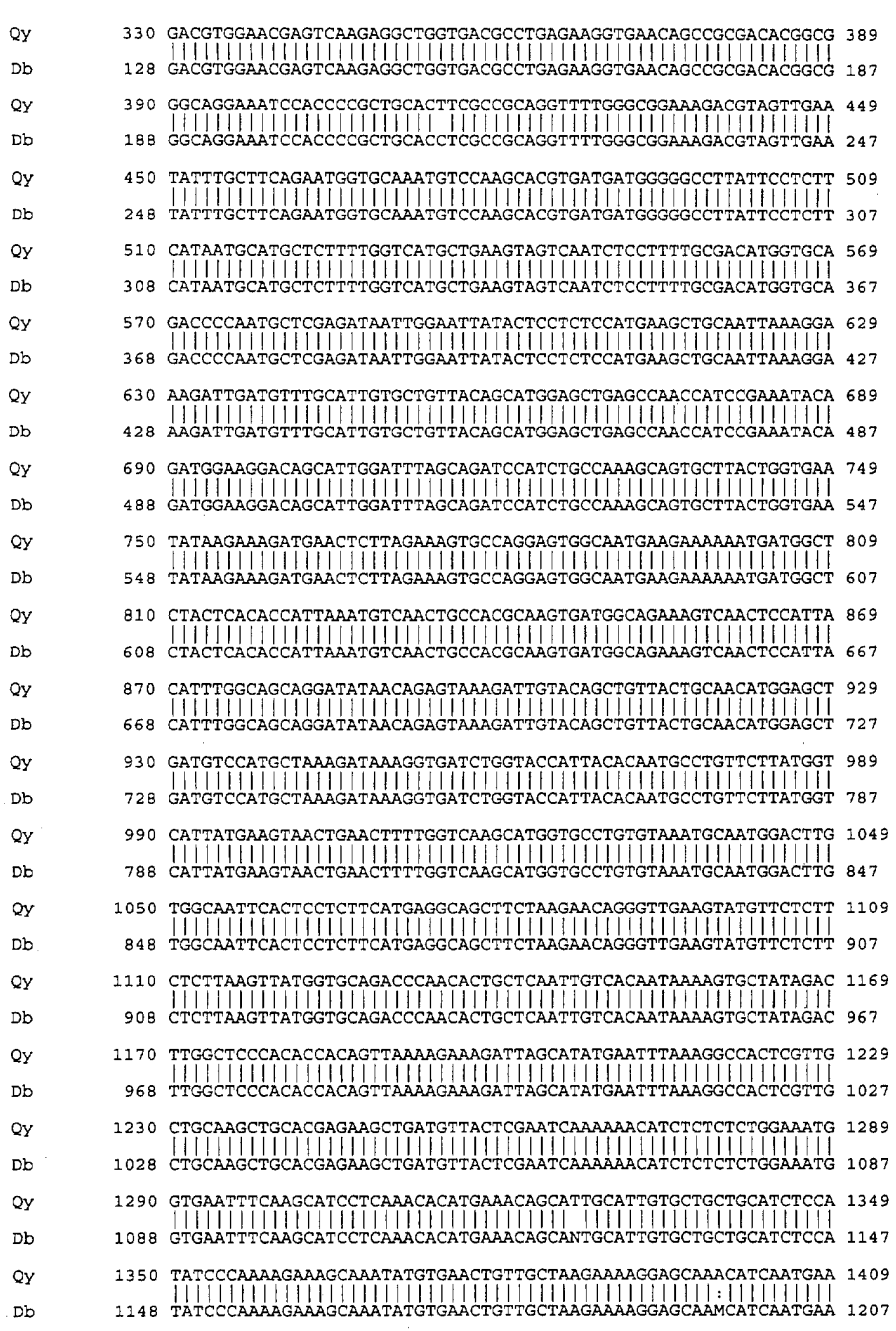
```
Db       2048 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2107
Qy       2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
Db       2108 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2167
Qy       2370 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2429
Db       2168 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2227
Qy       2430 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2489
Db       2228 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2287
Qy       2490 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2549
Db       2288 GGACGAACACAGCTTTGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2347
Qy       2550 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2609
Db       2348 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2407
Qy       2610 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2669
Db       2408 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2467
Qy       2670 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729
Db       2468 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2527
Qy       2730 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2789
Db       2528 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2587
Qy       2790 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2849
Db       2588 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2647
Qy       2850 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2909
Db       2648 TTTAGCATAACTCAATTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2707
Qy       2910 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2969
Db       2708 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2767
Qy       2970 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 3029
Db       2768 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 2827
Qy       3030 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 3089
Db       2828 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 2887
Qy       3090 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 3149
Db       2888 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 2947
Qy       3150 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3209
Db       2948 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3007
Qy       3210 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3269
Db       3008 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3067
Qy       3270 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3329
Db       3068 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3127
Qy       3330 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3389
Db       3128 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3187
```

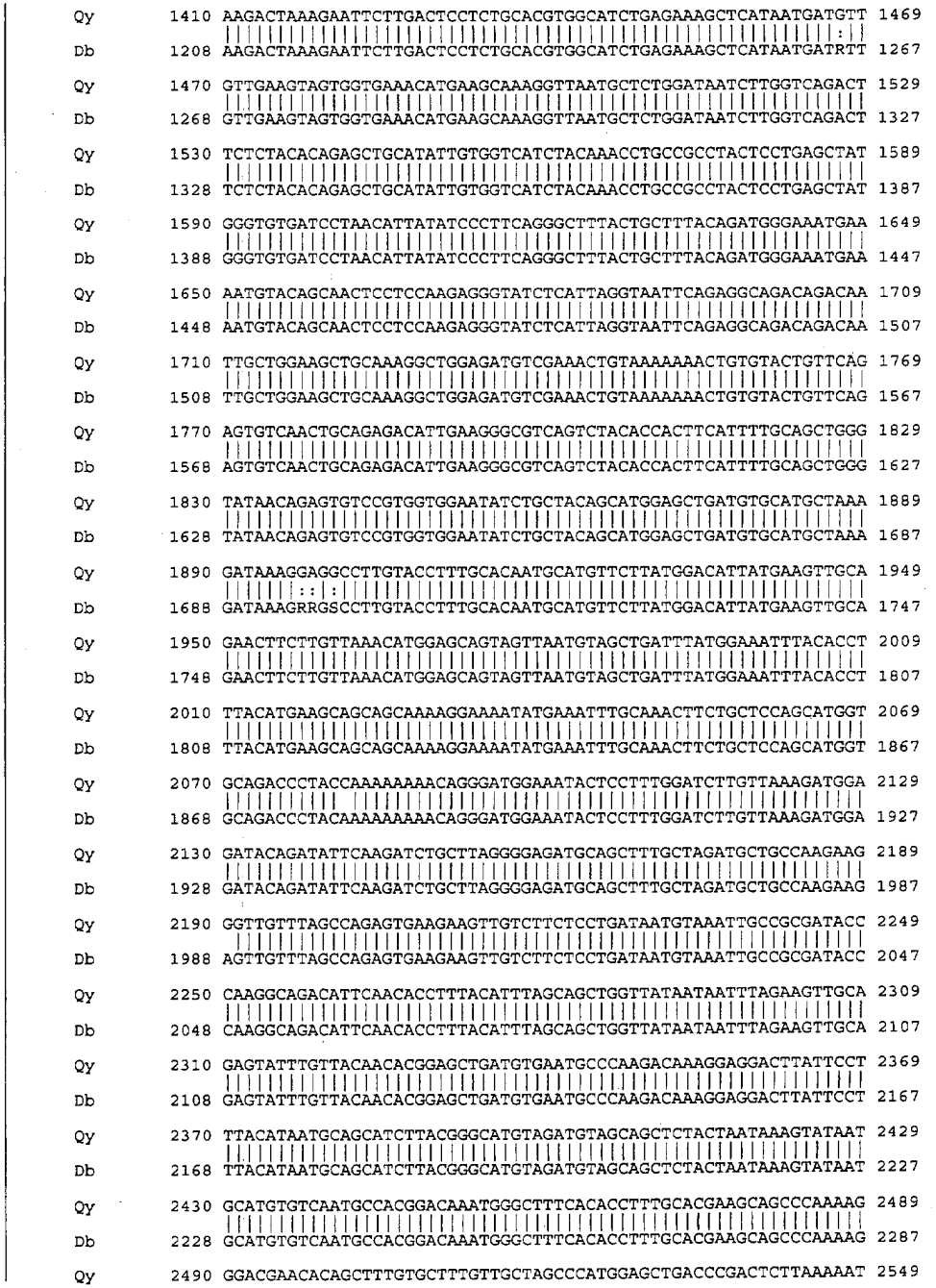
```
Qy       3390 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3449
Db       3188 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3247
Qy       3450 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3509
Db       3248 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3307
Qy       3510 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3569
Db       3308 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3367
Qy       3570 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3629
Db       3368 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3427
Qy       3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3689
Db       3428 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3487
Qy       3690 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3749
Db       3488 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3547
Qy       3750 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3809
Db       3548 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3607
Qy       3810 AAAA 3813
Db       3608 AAAA 3611
```

RESULT 13
AX076213
LOCUS AX076213 4512 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 3 from Patent WO0104326.
ACCESSION AX076213
VERSION AX076213.1 GI:12710838
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Berthelsen,J., Toma,S. and Isacchi,A.
TITLE Tankyrase homolog protein (thp), nucleic acids, and methods related to the same
JOURNAL Patent: WO 0104326-A 3 18-JAN-2001;
PHARMACIA & UPJOHN S.p.A. (IT)
FEATURES Location/Qualifiers
source 1. .4512
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
unsure 1124
/note="n is any base"
unsure 2672
/note="n is any base"
ORIGIN

Query Match 94.2%; Score 3594.8; DB 6; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3593; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```
Qy        210 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 269
Db          8 GCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGC 67
Qy        270 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 329
Db         68 GCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGG 127
```


Qy 330 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 389
Db 128 GACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCG 187
Qy 390 GGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 449
Db 188 GGCAGGAAATCCACCCCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAA 247
Qy 450 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 509
Db 248 TATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTT 307
Qy 510 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 569
Db 308 CATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCA 367
Qy 570 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 629
Db 368 GACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGA 427
Qy 630 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 689
Db 428 AAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACA 487
Qy 690 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 749
Db 488 GATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAA 547
Qy 750 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 809
Db 548 TATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCT 607
Qy 810 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 869
Db 608 CTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTA 667
Qy 870 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 929
Db 668 CATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCT 727
Qy 930 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 989
Db 728 GATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGT 787
Qy 990 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 1049
Db 788 CATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTG 847
Qy 1050 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 1109
Db 848 TGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTT 907
Qy 1110 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 1169
Db 908 CTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGAC 967
Qy 1170 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1229
Db 968 TTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTG 1027
Qy 1230 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1289
Db 1028 CTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATG 1087
Qy 1290 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCA 1349
Db 1088 GTGAATTTCAAGCATCCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCA 1147
Qy 1350 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAA 1409
Db 1148 TATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAA 1207


Qy 1410 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTT 1469
Db 1208 AAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTT 1267
Qy 1470 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1529
Db 1268 GTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACT 1327
Qy 1530 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1589
Db 1328 TCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTAT 1387
Qy 1590 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1649
Db 1388 GGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAA 1447
Qy 1650 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1709
Db 1448 AATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA 1507
Qy 1710 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1769
Db 1508 TTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAG 1567
Qy 1770 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1829
Db 1568 AGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGG 1627
Qy 1830 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1889
Db 1628 TATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAA 1687
Qy 1890 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1949
Db 1688 GATAAAGRRGSCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1747
Qy 1950 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 2009
Db 1748 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 1807
Qy 2010 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069
Db 1808 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 1867
Qy 2070 GCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 2129
Db 1868 GCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 1927
Qy 2130 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 2189
Db 1928 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 1987
Qy 2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
Db 1988 AGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2047
Qy 2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
Db 2048 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2107
Qy 2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
Db 2108 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2167
Qy 2370 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2429
Db 2168 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2227
Qy 2430 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2489
Db 2228 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2287
Qy 2490 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2549

```
Db       2288 GGACGAACACAGCTTTGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2347

Qy       2550 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2609

Db       2348 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2407

Qy       2610 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2669

Db       2408 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2467

Qy       2670 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729

Db       2468 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2527

Qy       2730 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2789

Db       2528 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2587

Qy       2790 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2849

Db       2588 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2647

Qy       2850 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2909

Db       2648 TTTAGCATAACTCAATTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2707

Qy       2910 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2969

Db       2708 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2767

Qy       2970 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 3029

Db       2768 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 2827

Qy       3030 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 3089

Db       2828 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 2887

Qy       3090 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 3149

Db       2888 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 2947

Qy       3150 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3209

Db       2948 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3007

Qy       3210 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3269

Db       3008 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3067

Qy       3270 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3329

Db       3068 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3127

Qy       3330 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3389

Db       3128 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3187

Qy       3390 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3449

Db       3188 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3247

Qy       3450 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3509

Db       3248 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3307

Qy       3510 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3569

Db       3308 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3367

Qy       3570 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3629

Db       3368 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3427

Qy       3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3689

Db       3428 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3487

Qy       3690 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3749

Db       3488 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3547

Qy       3750 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3809

Db       3548 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3607

Qy       3810 AAAA 3813

Db       3608 AAAA 3611
```

```
RESULT 14
AF264912
LOCUS       AF264912               5810 bp    mRNA    linear   PRI 14-MAY-2001
DEFINITION  Homo sapiens tankyrase-like protein (TNKL) mRNA, complete cds.
ACCESSION   AF264912
VERSION     AF264912.1  GI:12005975
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5810)
  AUTHORS   Kuimov,A.N., Kuprash,D.V., Petrov,V.N., Vdovichenko,K.K.,
            Scanlan,M.J., Jongeneel,C.V., Lagarkova,M.A. and Nedospasov,S.A.
  TITLE     Cloning and characterization of TNKL, a member of tankyrase gene
            family
  JOURNAL   Genes Immun. 2 (1), 52-55 (2001)
  MEDLINE   21190090
   PUBMED   11294570
REFERENCE   2  (bases 1 to 5810)
  AUTHORS   Kuimov,A.N., Kuprash,D.V., Petrov,V.N., Vdovichenko,K.K.,
            Kashkarova,U.A., Boitchenko,V.E., Nedospasov,S.A. and
            Lagarkova,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAY-2000) Laboratory of Molecular Immunology, A.N.
            Belozersky Institute of Physico-Chemical Biology, Moscow State
            University, Moscow 119899, Russia
FEATURES             Location/Qualifiers
     source          1. .5810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q22-10q24"
                     /tissue_type="breast carcinoma"
                     /note="derived from clones MO-BC-203 and I203"
     gene            <1. .5810
                     /gene="TNKL"
                     /note="a new member of tankyrase gene family; current
                     Unigene claster Hs.168491"
     CDS             1. .3501
                     /gene="TNKL"
                     /note="a new member of tankyrase family"
                     /codon_start=1
                     /product="tankyrase-like protein"
                     /protein_id="AAG44694.1"
                     /db_xref="GI:12005976"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
```

```
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
     STS             complement(1132. .1277)
                     /note="SHGC-53837"
     STS             complement(3665. .3975)
                     /note="WI-16054"
     STS             complement(5544. .5777)
                     /note="WI-8339"
ORIGIN

  Query Match          94.0%;  Score 3585.8;  DB 9;  Length 5810;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3587;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
Db        1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy      285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
Db       61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy      345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
Db      121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy      405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
Db      181 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy      465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
Db      241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy      525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
Db      301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy      585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
Db      361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

Qy      645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
Db      421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy      705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 764
Db      481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 540

Qy      765 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 824
Db      541 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 600

Qy      825 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
Db      601 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660

Qy      885 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 944
Db      661 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 720

Qy      945 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 1004
Db      721 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 780

Qy     1005 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 1064
Db      781 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 840

Qy     1065 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 1124
Db      841 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 900

Qy     1125 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 1184
Db      901 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 960

Qy     1185 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1244
Db      961 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1020

Qy     1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1304
Db     1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1080

Qy     1305 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364
Db     1081 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140

Qy     1365 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1424
Db     1141 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1200

Qy     1425 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484
Db     1201 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1260

Qy     1485 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544
Db     1261 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320

Qy     1545 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604
Db     1321 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380

Qy     1605 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1664
Db     1381 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1440

Qy     1665 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1724
Db     1441 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1500

Qy     1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
Db     1501 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1560

Qy     1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
Db     1561 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1620

Qy     1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
Db     1621 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1680

Qy     1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
Db     1681 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1740

Qy     1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
Db     1741 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 1800
```

```
Qy       2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2084
Db       1801 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAA 1860

Qy       2085 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2144
Db       1861 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 1920

Qy       2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2204
Db       1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 1980

Qy       2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
Db       1981 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2040

Qy       2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
Db       2041 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2100

Qy       2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
Db       2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2160

Qy       2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
Db       2161 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220

Qy       2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
Db       2221 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2280

Qy       2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
Db       2281 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2340

Qy       2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
Db       2341 CCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2400

Qy       2625 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
Db       2401 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2460

Qy       2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
Db       2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2520

Qy       2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
Db       2521 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2580

Qy       2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
Db       2581 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640

Qy       2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
Db       2641 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2700

Qy       2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
Db       2701 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2760

Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
Db       2761 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 2820

Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
Db       2821 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 2880

Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
Db       2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 2940

Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
Db       2941 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3000

Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
Db       3001 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3060

Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
Db       3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3120

Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
Db       3121 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3180

Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
Db       3181 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3240

Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
Db       3241 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3300

Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
Db       3301 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3360

Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
Db       3361 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3420

Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
Db       3421 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3480

Qy       3705 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3764
Db       3481 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3540

Qy       3765 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3541 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3589
```

```
RESULT 15
AX062142
LOCUS       AX062142               3508 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 1 from Patent WO0100849.
ACCESSION   AX062142
VERSION     AX062142.1  GI:12540045
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 1 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3508
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .>3508
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25052.1"
```

/db_xref="GI:12540046"
/db_xref="REMTREMBL:CAC25052"
/translation="ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVK
RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNAC
SFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG
RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPL
HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM
DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFK
GHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK
GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ
TCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV
ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPL
HNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKK
NRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRH
STPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNAC
VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLT
AAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSS
VVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKE
LKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE
EEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER
MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK
DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYR
GEQAYPEYLITYQIMRPEGMVDG"

ORIGIN

Query Match 91.8%; Score 3504.8; DB 6; Length 3508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3506; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600

Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840

Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900

Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960

Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020

Qy       1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080

Qy       1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140

Qy       1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200

Qy       1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260

Qy       1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320

Qy       1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380

Qy       1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440

Qy       1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500

Qy       1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560

Qy       1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1620

Qy       1835 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680

Qy       1895 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1740

Qy       1955 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 2014
```

```
Db      1741 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 1800
Qy      2015 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
Db      1801 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860
Qy      2075 CCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 2134
Db      1861 CCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 1920
Qy      2135 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
Db      1921 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980
Qy      2195 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2254
Db      1981 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2040
Qy      2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2314
Db      2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2100
Qy      2315 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
Db      2101 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160
Qy      2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2434
Db      2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2220
Qy      2435 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
Db      2221 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280
Qy      2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
Db      2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340
Qy      2555 AGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
Db      2341 AGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGC 2400
Qy      2615 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2674
Db      2401 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2460
Qy      2675 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2734
Db      2461 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2520
Qy      2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
Db      2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580
Qy      2795 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2854
Db      2581 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2640
Qy      2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2914
Db      2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2700
Qy      2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2974
Db      2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2760
Qy      2975 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 3034
Db      2761 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 2820
Qy      3035 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 3094
Db      2821 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 2880
Qy      3095 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 3154
Db      2881 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 2940
Qy      3155 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3214
Db      2941 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3000
Qy      3215 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3274
Db      3001 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3060
Qy      3275 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3334
Db      3061 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3120
Qy      3335 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3394
Db      3121 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3180
Qy      3395 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3454
Db      3181 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3240
Qy      3455 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3514
Db      3241 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3300
Qy      3515 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3574
Db      3301 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3360
Qy      3575 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3634
Db      3361 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3420
Qy      3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3694
Db      3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3480
Qy      3695 TATGAGGCCTGAAGGTATGGTCGATGGA 3722
Db      3481 TATGAGGCCTGAAGGTATGGTCGATGGA 3508
```

```
RESULT 16
AF342982
LOCUS       AF342982               3501 bp    mRNA    linear   PRI 18-SEP-2001
DEFINITION  Homo sapiens tankyrase 2 mRNA, complete cds.
ACCESSION   AF342982
VERSION     AF342982.1  GI:13430364
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3501)
  AUTHORS   Kaminker,P.G., Kim,S.H., Taylor,R.D., Zebarjadian,Y., Funk,W.D.,
            Morin,G.B., Yaswen,P. and Campisi,J.
  TITLE     TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
            rapid induction of cell death upon overexpression
  JOURNAL   J. Biol. Chem. 276 (38), 35891-35899 (2001)
  MEDLINE   21443728
   PUBMED   11454873
REFERENCE   2  (bases 1 to 3501)
  AUTHORS   Kaminker,P.G., Campisi,J., Kim,S.H., Yaswen,P. and Morin,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2001) Life Sciences, Lawrence Berkeley National
            Labs, 1 Cyclotron Rd 84-171, Berkeley, CA 94720, USA
FEATURES             Location/Qualifiers
     source          1. .3501
```

```
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q23-q24"
     CDS             1. .3501
                     /note="TNKS2; telomere-associated poly(ADP-ribose)
                     polymerase; contains ankyrin domain, sterile alpha motif
                     and PARP catalytic domain; similar to TNKS"
                     /codon_start=1
                     /product="tankyrase 2"
                     /protein_id="AAK25811.1"
                     /db_xref="GI:13430365"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAARSADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match           91.7%;  Score 3499;  DB 9;  Length 3501;
  Best Local Similarity 99.9%;  Pred. No. 0;
  Matches 3499;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
Db         1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy       285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
Db        61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy       345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
Db       121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy       405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
Db       181 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy       465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
Db       241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy       525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
Db       301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy       585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
Db       361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

Qy       645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
Db       421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy       705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 764
Db       481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 540

Qy       765 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 824
Db       541 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 600

Qy       825 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
Db       601 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660

Qy       885 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 944
Db       661 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 720

Qy       945 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 1004
Db       721 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 780

Qy      1005 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 1064
Db       781 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 840

Qy      1065 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 1124
Db       841 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 900

Qy      1125 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 1184
Db       901 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 960

Qy      1185 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1244
Db       961 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1020

Qy      1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1304
Db      1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1080

Qy      1305 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364
Db      1081 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140

Qy      1365 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1424
Db      1141 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1200

Qy      1425 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484
Db      1201 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1260

Qy      1485 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544
Db      1261 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320

Qy      1545 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604
Db      1321 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380

Qy      1605 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1664
Db      1381 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1440

Qy      1665 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1724
Db      1441 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1500

Qy      1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
Db      1501 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1560

Qy      1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
```

```
Db       1561 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1620
Qy       1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
Db       1621 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1680
Qy       1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
Db       1681 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1740
Qy       1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
Db       1741 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 1800
Qy       2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2084
Db       1801 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAA 1860
Qy       2085 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2144
Db       1861 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 1920
Qy       2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2204
Db       1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 1980
Qy       2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
Db       1981 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2040
Qy       2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
Db       2041 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2100
Qy       2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
Db       2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2160
Qy       2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
Db       2161 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220
Qy       2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
Db       2221 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2280
Qy       2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
Db       2281 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2340
Qy       2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
Db       2341 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2400
Qy       2625 TCTGCTCTGCCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
Db       2401 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2460
Qy       2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
Db       2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2520
Qy       2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
Db       2521 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2580
Qy       2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
Db       2581 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640
Qy       2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
Db       2641 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2700
```

```
Qy       2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
Db       2701 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGWATCAATGCTTAT 2760
Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
Db       2761 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 2820
Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
Db       2821 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 2880
Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
Db       2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 2940
Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
Db       2941 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3000
Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
Db       3001 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3060
Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
Db       3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3120
Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
Db       3121 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3180
Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
Db       3181 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3240
Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
Db       3241 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3300
Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
Db       3301 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3360
Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
Db       3361 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3420
Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
Db       3421 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3480
Qy       3705 GAAGGTATGGTCGATGGATAA 3725
Db       3481 GAAGGTATGGTCGATGGATAA 3501
```

```
RESULT 17
AR232162
LOCUS       AR232162               3498 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 4 from patent US 6455290.
ACCESSION   AR232162
VERSION     AR232162.1  GI:27274049
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3498)
  AUTHORS   Berthelsen,J., Toma,S. and Isacchi,A.
  TITLE     Tankyrase homolog protein (THP), nucleic acids, and methods related
            to the same
  JOURNAL   Patent: US 6455290-A 4 24-SEP-2002;
```

```
FEATURES             Location/Qualifiers
     source          1. .3498
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             91.4%;  Score 3488.8;  DB 6;  Length 3498;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3487;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy        285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy        345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy        405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy        465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy        525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy        585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

Qy        645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy        705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 540

Qy        765 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 600

Qy        825 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660

Qy        885 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 720

Qy        945 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 780

Qy       1005 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 840

Qy       1065 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 900

Qy       1125 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 960

Qy       1185 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1020

Qy       1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1080

Qy       1305 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        1081 CCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140

Qy       1365 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1424
              |||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        1141 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTC 1200

Qy       1425 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484
              |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        1201 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTG 1260

Qy       1485 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320

Qy       1545 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380

Qy       1605 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1440

Qy       1665 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1500

Qy       1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1560

Qy       1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1620

Qy       1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||::|:||||
Db        1621 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTT 1680

Qy       1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1740

Qy       1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 1800

Qy       2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2084
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        1801 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAA 1860

Qy       2085 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 1920

Qy       2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2204
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGA 1980

Qy       2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2040
```

```
Qy       2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2100

Qy       2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2160

Qy       2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220

Qy       2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2280

Qy       2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
             |||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2340

Qy       2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2400

Qy       2625 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2460

Qy       2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2520

Qy       2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2580

Qy       2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640

Qy       2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2700

Qy       2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2760

Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 2820

Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 2880

Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 2940

Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3000

Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3060

Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3120

Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3180

Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3240

Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3300

Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3360

Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3420

Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3480

Qy       3705 GAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||
Db       3481 GAAGGTATGGTCGATGGA 3498
```

RESULT 18
AX076214

```
LOCUS       AX076214               3498 bp    DNA     linear   PAT 06-FEB-2001
DEFINITION  Sequence 4 from Patent WO0104326.
ACCESSION   AX076214
VERSION     AX076214.1  GI:12710839
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Berthelsen,J., Toma,S. and Isacchi,A.
  TITLE     Tankyrase homolog protein (thp), nucleic acids, and methods related
            to the same
  JOURNAL   Patent: WO 0104326-A 4 18-JAN-2001;
            PHARMACIA & UPJOHN S.p.A. (IT)
FEATURES             Location/Qualifiers
     source          1. .3498
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .>3498
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC28482.1"
                     /db_xref="GI:12710840"
                     /db_xref="REMTREMBL:CAC28482"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHLAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETAXHCAAASPYPKRKQICELLLRKGAX
                     INEKTKEFLTPLHVASEKAHNDXVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKXXLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKSCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALXLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRXLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
```

```
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
     unsure          368
                     /note="Xaa is Met or Leu or Val"
     unsure          392
                     /note="Xaa is Asn or Thr"
     unsure          415
                     /note="Xaa is Val or Ile"
     unsure          558
                     /note="Xaa is Gly or Glu"
     unsure          559
                     /note="Xaa is Gly or Ala"
     unsure          764
                     /note="Xaa is Leu or Val"
     unsure          884
                     /note="Xaa is Asn or His or Asp or Tyr"
ORIGIN

  Query Match             91.4%;  Score 3488.8;  DB 6;  Length 3498;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3487;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy        285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy        345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy        405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCTGCACCTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy        465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy        525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy        585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

Qy        645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy        705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 540

Qy        765 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 600

Qy        825 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660

Qy        885 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 720

Qy        945 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 1004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 780

Qy       1005 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 1064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 840

Qy       1065 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 1124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 900

Qy       1125 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 1184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 960

Qy       1185 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1020

Qy       1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1080

Qy       1305 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTCAAACACATGAAACAGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140

Qy       1365 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1424
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db       1141 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTC 1200

Qy       1425 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db       1201 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTG 1260

Qy       1485 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320

Qy       1545 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380

Qy       1605 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1440

Qy       1665 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1500

Qy       1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1560

Qy       1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1620

Qy       1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||::|:||||
Db       1621 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTT 1680

Qy       1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1740

Qy       1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 1800

Qy       2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2084
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
```

```
Db       1801 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAA 1860
Qy       2085 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2144
Db       1861 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 1920
Qy       2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2204
Db       1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGA 1980
Qy       2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
Db       1981 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2040
Qy       2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
Db       2041 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2100
Qy       2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
Db       2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2160
Qy       2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
Db       2161 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220
Qy       2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
Db       2221 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2280
Qy       2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
Db       2281 TGTGCTTTGKTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2340
Qy       2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
Db       2341 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2400
Qy       2625 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
Db       2401 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2460
Qy       2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
Db       2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2520
Qy       2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
Db       2521 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2580
Qy       2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
Db       2581 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640
Qy       2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
Db       2641 TTCGTAAGGNATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2700
Qy       2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
Db       2701 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2760
Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
Db       2761 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 2820
Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
Db       2821 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 2880
Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
Db       2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 2940
Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
Db       2941 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3000
Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
Db       3001 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3060
Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
Db       3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3120
Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
Db       3121 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3180
Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
Db       3181 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3240
Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
Db       3241 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3300
Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
Db       3301 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3360
Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
Db       3361 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3420
Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
Db       3421 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3480
Qy       3705 GAAGGTATGGTCGATGGA 3722
Db       3481 GAAGGTATGGTCGATGGA 3498
```

RESULT 19
AR392872

```
LOCUS       AR392872               3797 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1 from patent US 6617102.
ACCESSION   AR392872
VERSION     AR392872.1  GI:40118097
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3797)
  AUTHORS   Luo,Y., Chan,E., Xu,X., Huang,B. and Ossovskaya,V.
  TITLE     Tankyrase H, compositions involved in the cell cycle and methods of
            use
  JOURNAL   Patent: US 6617102-A 1 09-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .3797
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match 88.9%; Score 3393.4; DB 6; Length 3797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481
Db        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Qy        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541
```

```
Db        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Qy        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601
Db        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Qy        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661
Db        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Qy        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721
Db        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Qy        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781
Db        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Qy        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
Db        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Qy        842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
Db        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Qy        902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
Db        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Qy        962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
Db        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Qy       1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
Db       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Qy       1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
Db       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Qy       1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
Db       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Qy       1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
Db       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Qy       1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321
Db       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Qy       1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381
Db       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Qy       1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Db       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Qy       1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Db       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Qy       1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Db       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Qy       1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Db       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602

Qy       1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Db       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Qy       1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Db       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Qy       1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Db       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Qy       1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Db       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Qy       1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Db       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Qy       1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Db       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Qy       1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Db       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Qy       2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2101
Db       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Qy       2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Db       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Qy       2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Db       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Qy       2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Db       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Qy       2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Db       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Qy       2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Db       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Qy       2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Db       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Qy       2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Db       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Qy       2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Db       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Qy       2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Db       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Qy       2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Db       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
```

```
Qy       2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742

Qy       2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802

Qy       2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862

Qy       2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922

Qy       2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982

Qy       3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042

Qy       3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102

Qy       3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162

Qy       3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

Qy       3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

Qy       3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342

Qy       3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402

Qy       3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462

Qy       3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

Qy       3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Qy       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Qy       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Qy       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Qy       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
```

```
             ||||||||||||||||||||||||||||||||||||
Db       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
```

```
RESULT 20
AR352002
LOCUS       AR352002               3394 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 1 from patent US 6589725.
ACCESSION   AR352002
VERSION     AR352002.1  GI:33756953
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3394)
  AUTHORS   Luo,Y., Chan,E., Xu,X. and Huang,B.
  TITLE     Tankyrase H, compositions involved in the cell cycle and methods of
            use
  JOURNAL   Patent: US 6589725-A 1 08-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .3394
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           88.8%;  Score 3387.6;  DB 6;  Length 3394;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3390;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy        423 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 60

Qy        483 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 120

Qy        543 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 180

Qy        603 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 240

Qy        663 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 722
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGCTGAGCCAACCATCCTAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 300

Qy        723 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 360

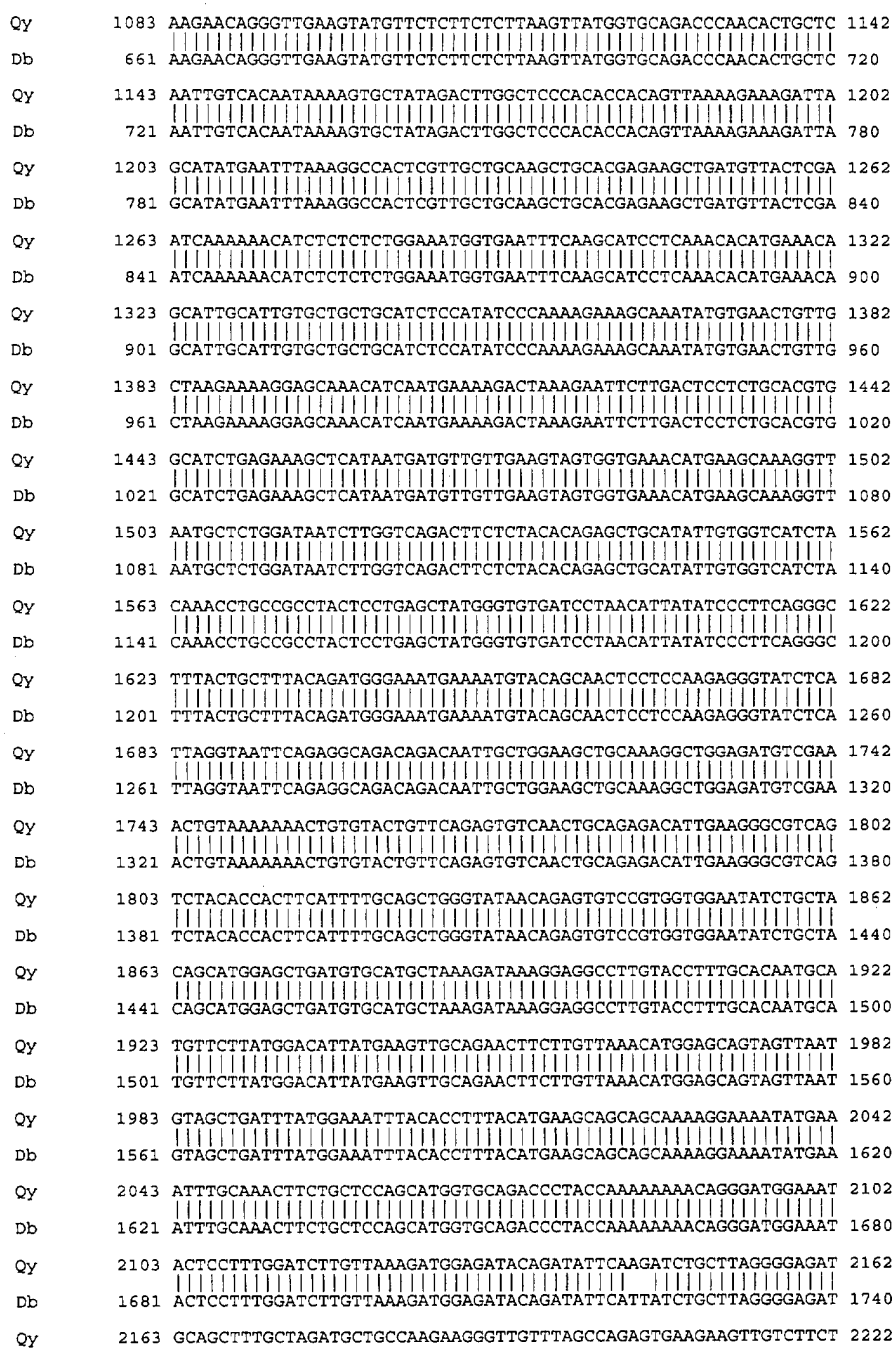
Qy        783 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 420

Qy        843 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480

Qy        903 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 540

Qy        963 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 600

Qy       1023 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 660
```


Qy 1083 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1142
Db 661 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 720
Qy 1143 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1202
Db 721 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 780
Qy 1203 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1262
Db 781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 840
Qy 1263 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1322
Db 841 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 900
Qy 1323 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1382
Db 901 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 960
Qy 1383 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1442
Db 961 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1020
Qy 1443 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1502
Db 1021 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1080
Qy 1503 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1562
Db 1081 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1140
Qy 1563 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1622
Db 1141 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1200
Qy 1623 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1682
Db 1201 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1260
Qy 1683 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1742
Db 1261 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1320
Qy 1743 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1802
Db 1321 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1380
Qy 1803 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1862
Db 1381 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1440
Qy 1863 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1922
Db 1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1500
Qy 1923 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1982
Db 1501 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1560
Qy 1983 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2042
Db 1561 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620
Qy 2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 2102
Db 1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 1680
Qy 2103 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2162
Db 1681 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCATTATCTGCTTAGGGGAGAT 1740
Qy 2163 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2222
Db 1741 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 1800
Qy 2223 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2282
Db 1801 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 1860
Qy 2283 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2342
Db 1861 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 1920
Qy 2343 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2402
Db 1921 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 1980
Qy 2403 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2462
Db 1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2040
Qy 2463 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2522
Db 2041 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2100
Qy 2523 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2582
Db 2101 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160
Qy 2583 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2642
Db 2161 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2220
Qy 2643 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2702
Db 2221 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2280
Qy 2703 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2762
Db 2281 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2340
Qy 2763 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2822
Db 2341 AGTTTTTCAGAACTGTCTTCATTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2400
Qy 2823 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2882
Db 2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2460
Qy 2883 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2942
Db 2461 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2520
Qy 2943 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 3002
Db 2521 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2580
Qy 3003 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 3062
Db 2581 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 2640
Qy 3063 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 3122
Db 2641 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 2700
Qy 3123 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 3182
Db 2701 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 2760
Qy 3183 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3242
Db 2761 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 2820
Qy 3243 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3302

```
Db       2821 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 2880

Qy       3303 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 2940

Qy       3363 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3000

Qy       3423 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3060

Qy       3483 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3120

Qy       3543 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3180

Qy       3603 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3240

Qy       3663 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3300

Qy       3723 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3360

Qy       3783 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
             ||||||||||||||||||||||||||||||||||
Db       3361 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAA 3394
```

RESULT 21
AX133694
LOCUS AX133694 3394 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0130987.
ACCESSION AX133694
VERSION AX133694.1 GI:14139712
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Luo,Y., Chan,E., Xu,X. and Huang,B.
TITLE Tankyrase h, compositions involved in the cell cycle and methods of use
JOURNAL Patent: WO 0130987-A 1 03-MAY-2001;
Rigel Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1. .3394
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 88.8%; Score 3387.6; DB 6; Length 3394;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy        423 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 60

Qy        483 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          61 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 120

Qy        543 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 180

Qy        603 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 240

Qy        663 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGCTGAGCCAACCATCCTAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 300

Qy        723 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 360

Qy        783 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 420

Qy        843 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480

Qy        903 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 540

Qy        963 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 600

Qy       1023 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 660

Qy       1083 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 720

Qy       1143 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 780

Qy       1203 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 840

Qy       1263 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 900

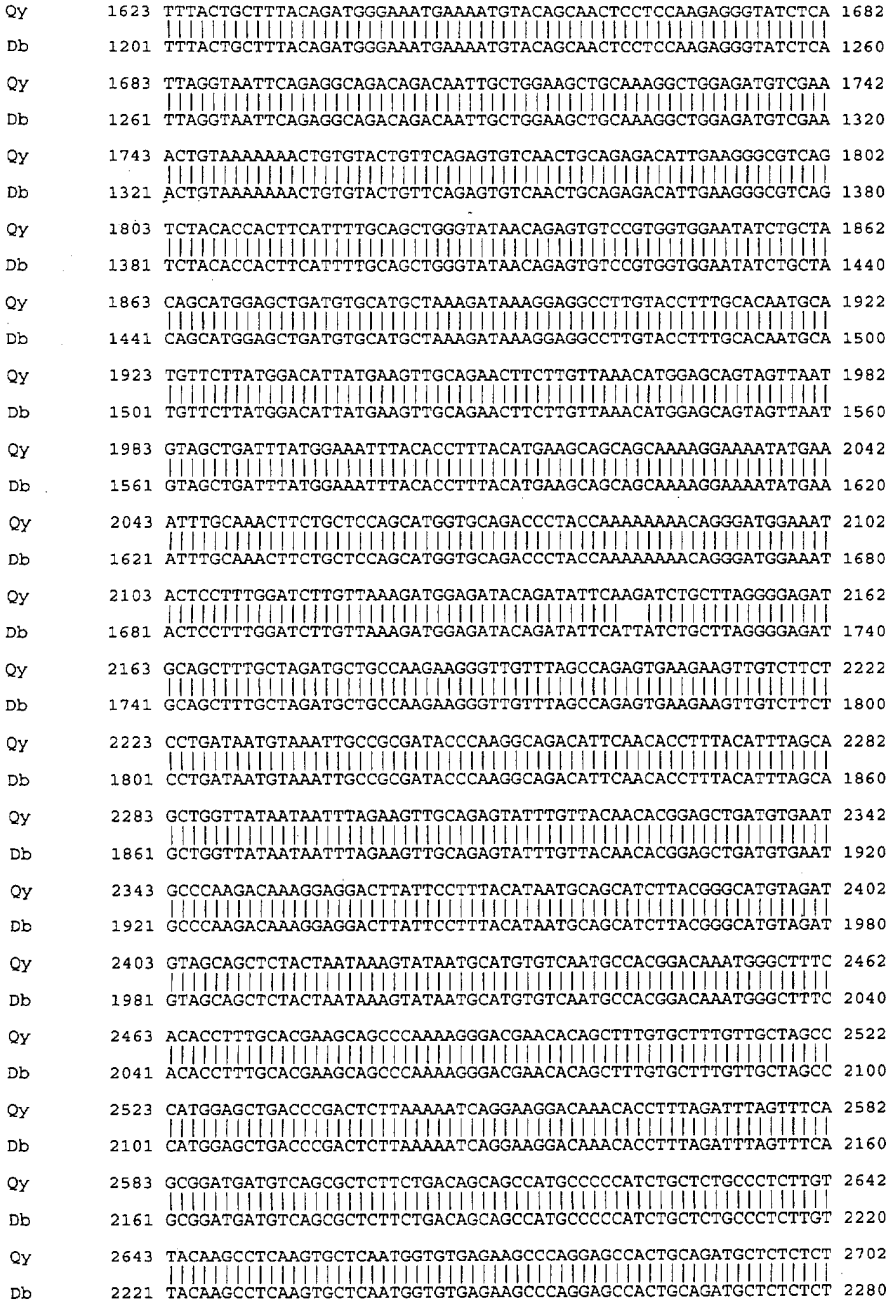
Qy       1323 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 960

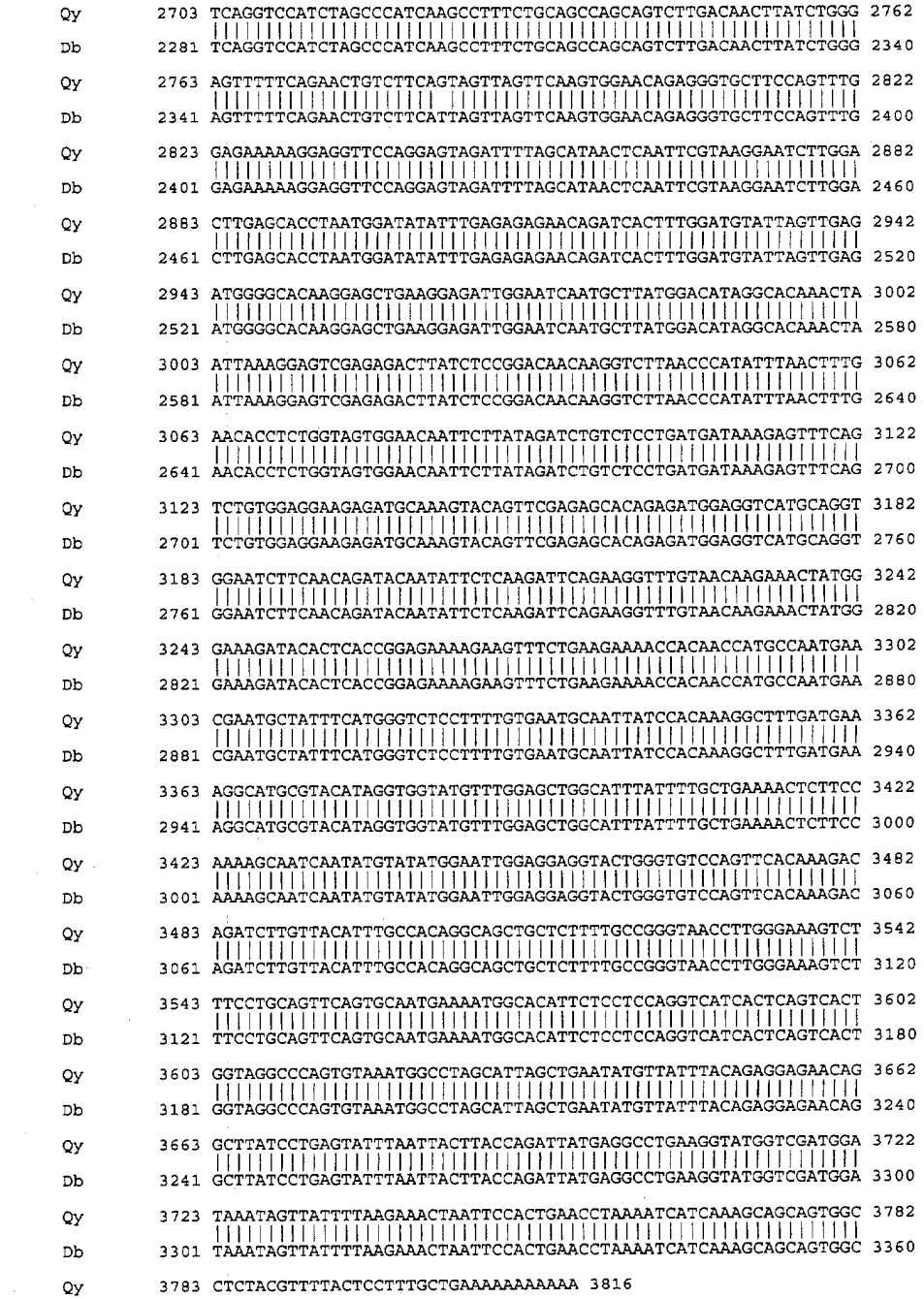
Qy       1383 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1020

Qy       1443 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1080

Qy       1503 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1140

Qy       1563 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1200
```


Qy 1623 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1682
Db 1201 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1260
Qy 1683 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1742
Db 1261 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1320
Qy 1743 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1802
Db 1321 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1380
Qy 1803 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1862
Db 1381 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1440
Qy 1863 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1922
Db 1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1500
Qy 1923 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1982
Db 1501 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1560
Qy 1983 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2042
Db 1561 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620
Qy 2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 2102
Db 1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 1680
Qy 2103 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2162
Db 1681 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCATTATCTGCTTAGGGGAGAT 1740
Qy 2163 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2222
Db 1741 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 1800
Qy 2223 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2282
Db 1801 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 1860
Qy 2283 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2342
Db 1861 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 1920
Qy 2343 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2402
Db 1921 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 1980
Qy 2403 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2462
Db 1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2040
Qy 2463 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2522
Db 2041 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2100
Qy 2523 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2582
Db 2101 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160
Qy 2583 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2642
Db 2161 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2220
Qy 2643 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2702
Db 2221 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2280


Qy 2703 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2762
Db 2281 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2340
Qy 2763 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2822
Db 2341 AGTTTTTCAGAACTGTCTTCATTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2400
Qy 2823 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2882
Db 2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2460
Qy 2883 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2942
Db 2461 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2520
Qy 2943 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 3002
Db 2521 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2580
Qy 3003 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 3062
Db 2581 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 2640
Qy 3063 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 3122
Db 2641 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 2700
Qy 3123 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 3182
Db 2701 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 2760
Qy 3183 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3242
Db 2761 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 2820
Qy 3243 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3302
Db 2821 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 2880
Qy 3303 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3362
Db 2881 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 2940
Qy 3363 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3422
Db 2941 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3000
Qy 3423 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3482
Db 3001 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3060
Qy 3483 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3542
Db 3061 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3120
Qy 3543 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3602
Db 3121 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3180
Qy 3603 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
Db 3181 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3240
Qy 3663 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
Db 3241 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3300
Qy 3723 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3782
Db 3301 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3360
Qy 3783 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816

```
Db       3361 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3394


RESULT 22
AR361485
LOCUS       AR361485                4297 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 3 from patent US 6599728.
ACCESSION   AR361485
VERSION     AR361485.1  GI:33769333
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4297)
  AUTHORS   Morin,G.B., Funk,W.D. and Piatyszek,M.A.
  TITLE     Second mammalian tankyrase
  JOURNAL   Patent: US 6599728-A 3 29-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .4297
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             87.6%;  Score 3341;  DB 6;  Length 4297;
  Best Local Similarity   89.0%;  Pred. No. 0;
  Matches 3400;  Conservative    8;  Mismatches  402;  Indels   12;  Gaps    2;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| |  |||||||||||||||||||||||||||||||||||| ||| ||
Db         74 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAANGGGCGCGC 133

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
                ||    | |||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTGGGCGCGGCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 193

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 253

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        254 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCANGATCATGTCGGGTCGCCGCT 313

Qy        241 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              |||||||||||||| |||||||||||| |||||||||||| |||||||| ||||||||||
Db        314 GCGCCGGCGGGGGGANCGGCCTGCGCGANCGCCGCGGCCGAAGCCGTGGAACCGGCCGCCC 373

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              || | |||||||| |||||||||||||||||||||||||||||| |||||| ||||||||
Db        374 GAAANCTGTTCGAAGCGTGCCGCAACGGGGACGTGGAACGANTCAAGAAGCTGGTGACNC 433

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||:|||||||||||||||||||||| |||||||||||||||||||||||||||: |||
Db        434 CTGARAAGGTGAACAGCCGCGACACNGCGGGCAGGAAATCCACCCCGCTGCACTTYCCCG 493

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              || |||||||||||||||| || ||:|||||||||||||| ||||||||||||||:|||
Db        494 CANGTTTTGGGCGGAAAGACTTANTTRAATATTTGCTTCANAATGGTGCAAATGTYCAAN 553

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              |||||||| ||||||||||||||||||||||||||||||||||||:||||:||
Db        554 CACGTGATNATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCMTGCTRAAA 613

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
                || |||||||||||| |||| ||||| |:|||||||||||||||||||||||||||||
Db        614 NTATCNATCTCCTTTTGCNACATNGTGCANAMCCCAATGCTCGAGATAATTGGAATTATA 673

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              |||||| | |||||||||||||||||||||||||||||||  ||||||||||||||||||
Db        674 CTCCTCNCNATGAAGCTGCAATTAAAGGAAAGATTGANNNTTGCATTGTGCTGTTACAGC 733

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 793

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 853

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 913

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 973

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1033

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1093

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1153

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1213

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1273

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1333

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1393

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1453

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1513

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1573

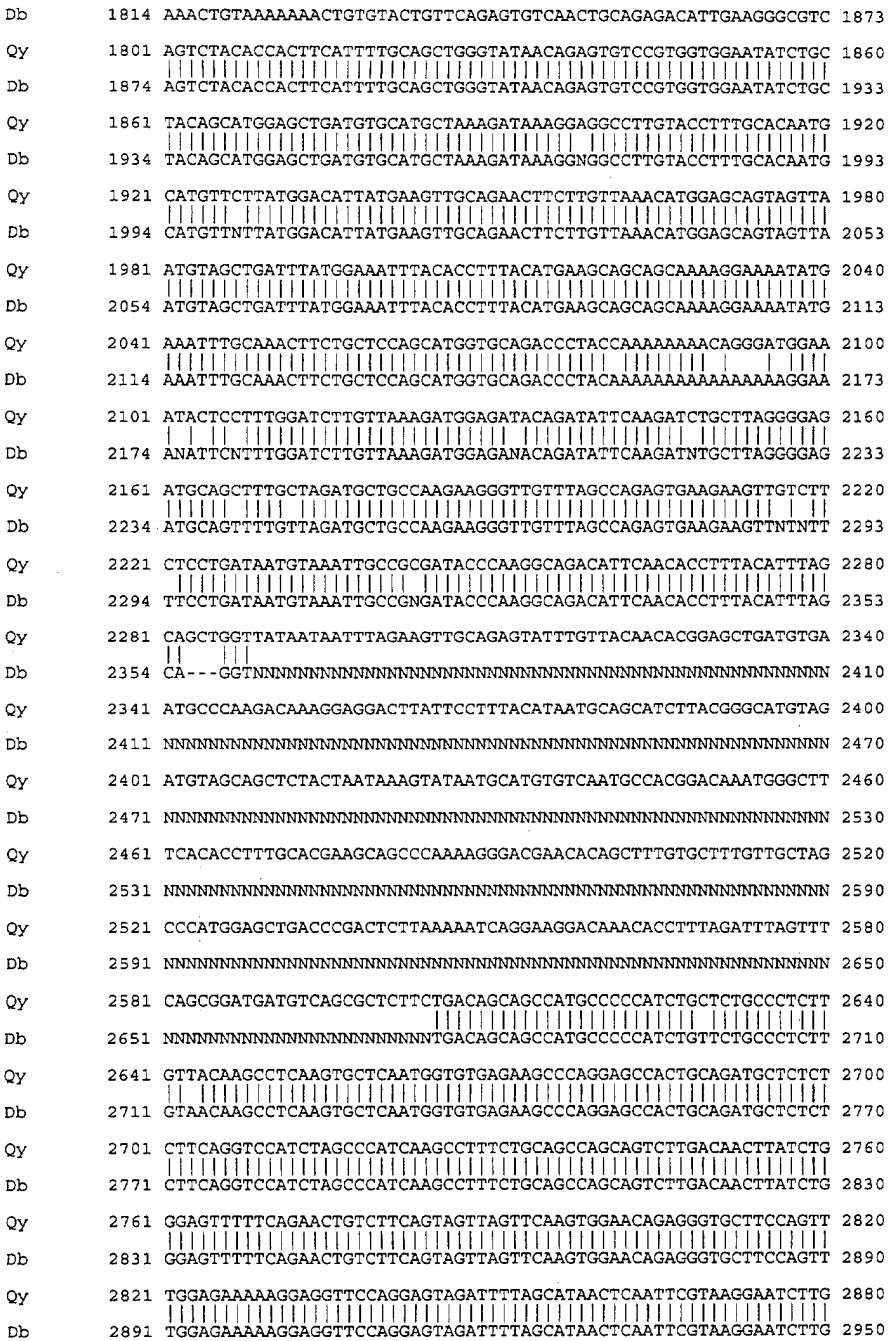
Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1633

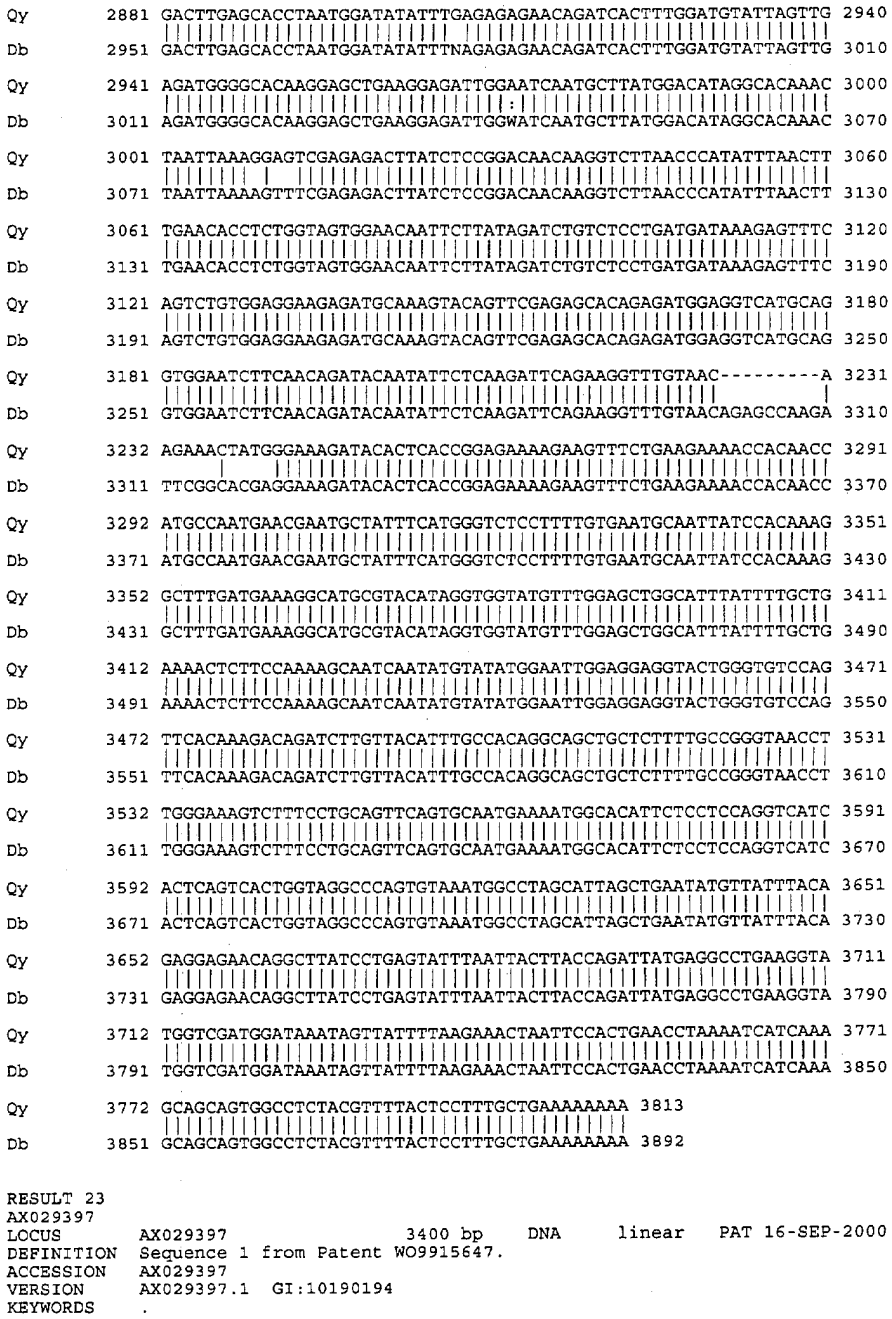
Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1634 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1693

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1694 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1753

Qy       1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1754 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1813

Qy       1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```





RESULT 23
AX029397
LOCUS AX029397 3400 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 1 from Patent WO9915647.
ACCESSION AX029397
VERSION AX029397.1 GI:10190194
KEYWORDS .

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Sutherland,R.L. and Daly,R.J.
  TITLE     A potential effector for the grb7 family of signalling proteins
  JOURNAL   Patent: WO 9915647-A 1 01-APR-1999;
            GARVAN INST MED RES (AU) ; SUTHERLAND ROBERT LYNDSAY (AU) ; DALY
            ROGER JOHN (AU)
FEATURES             Location/Qualifiers
     source          1. .3400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           86.4%;  Score 3297;  DB 6;  Length 3400;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        501 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60

Qy        561 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120

Qy        621 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180

Qy        681 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240

Qy        741 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300

Qy        801 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360

Qy        861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420

Qy        921 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 980
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480

Qy        981 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1040
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540

Qy       1041 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600

Qy       1101 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1160
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660

Qy       1161 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720

Qy       1221 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780

Qy       1281 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840

Qy       1341 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900

Qy       1401 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960

Qy       1461 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020

Qy       1521 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080

Qy       1581 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy       1641 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200

Qy       1701 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260

Qy       1761 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320

Qy       1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380

Qy       1881 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440

Qy       1941 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500

Qy       2001 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560

Qy       2061 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2120
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620

Qy       2121 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680

Qy       2181 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740

Qy       2241 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800

Qy       2301 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860

Qy       2361 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420
```

```
Db       1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920

Qy       2421 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2480
Db       1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980

Qy       2481 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2540
Db       1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040

Qy       2541 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2600
Db       2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100

Qy       2601 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2660
Db       2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160

Qy       2661 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2720
Db       2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220

Qy       2721 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2780
Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280

Qy       2781 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2840
Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340

Qy       2841 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2900
Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400

Qy       2901 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2960
Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460

Qy       2961 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3020
Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520

Qy       3021 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3080
Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580

Qy       3081 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3140
Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640

Qy       3141 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3200
Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700

Qy       3201 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3260
Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3261 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3320
Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820

Qy       3321 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3380
Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880

Qy       3381 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3440
Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940

Qy       3441 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3500
```

```
Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000

Qy       3501 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3560
Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060

Qy       3561 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3620
Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120

Qy       3621 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3680
Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180

Qy       3681 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3740
Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240

Qy       3741 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3800
Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300

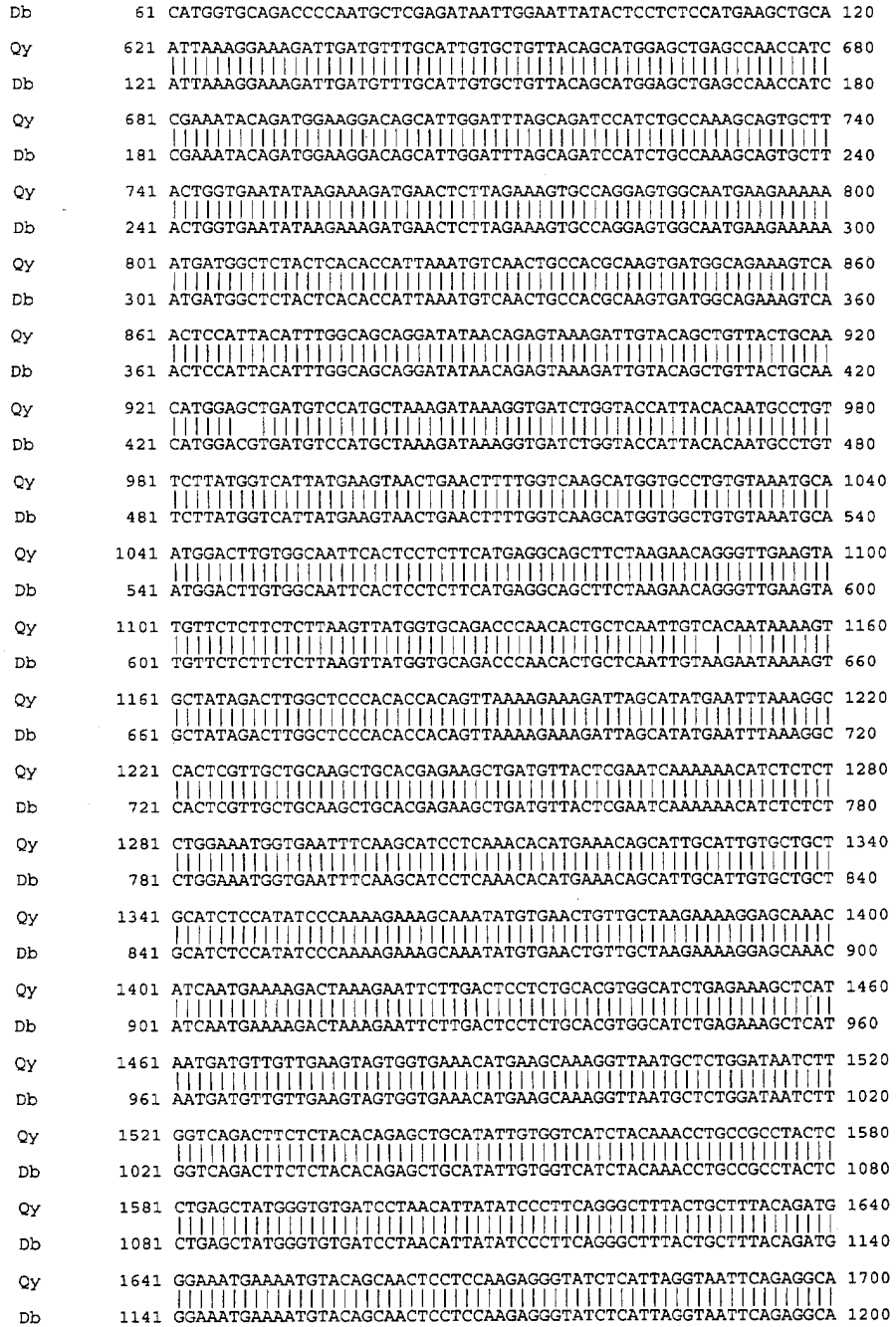
Qy       3801 TGCTGAAAAAAAA 3813
Db       3301 TGCTGAAAAAAAA 3313
```

```
RESULT 24
BD075726
LOCUS       BD075726               3400 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Potential effector of signal transducing protein GRB7 family.
ACCESSION   BD075726
VERSION     BD075726.1  GI:22621329
KEYWORDS    JP 2001517435-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3400)
  AUTHORS   Dary,R.J. and Sutherland,R.L.
  TITLE     Potential effector of signal transducing protein GRB7 family
  JOURNAL   Patent: JP 2001517435-A 1 09-OCT-2001;
            GARVAN INSTITUTE OF MEDICAL RESEARCH
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001517435-A/1
            PD   09-OCT-2001
            PF   23-SEP-1998 JP 2000512939
            PR   23-SEP-1997 AU      PO 9388
            PI   ROGER JOHN DARY,ROBERT LINDSEY SUTHERLAND
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12Q1/68,C12N15/00,C12N5/00
            CC   Potential effector of signal transducing protein GRB7 family
            FH   Key             Location/Qualifiers
            FT   source          1. .3400
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .3400
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           86.4%;  Score 3297;  DB 6;  Length 3400;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        501 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 560
Db          1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60

Qy        561 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 620
```


Db 61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120
Qy 621 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 680
Db 121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180
Qy 681 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 740
Db 181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240
Qy 741 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 800
Db 241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300
Qy 801 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 860
Db 301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360
Qy 861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 920
Db 361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420
Qy 921 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 980
Db 421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480
Qy 981 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1040
Db 481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540
Qy 1041 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1100
Db 541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600
Qy 1101 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1160
Db 601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660
Qy 1161 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1220
Db 661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720
Qy 1221 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1280
Db 721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780
Qy 1281 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1340
Db 781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840
Qy 1341 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1400
Db 841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900
Qy 1401 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1460
Db 901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960
Qy 1461 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1520
Db 961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020
Qy 1521 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1580
Db 1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080
Qy 1581 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1640
Db 1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140
Qy 1641 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1700
Db 1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200
Qy 1701 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1760
Db 1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260
Qy 1761 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1820
Db 1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320
Qy 1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1880
Db 1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380
Qy 1881 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1940
Db 1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440
Qy 1941 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 2000
Db 1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500
Qy 2001 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2060
Db 1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560
Qy 2061 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2120
Db 1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620
Qy 2121 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2180
Db 1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680
Qy 2181 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2240
Db 1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740
Qy 2241 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2300
Db 1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800
Qy 2301 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2360
Db 1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860
Qy 2361 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420
Db 1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920
Qy 2421 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2480
Db 1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980
Qy 2481 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2540
Db 1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040
Qy 2541 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2600
Db 2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100
Qy 2601 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2660
Db 2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160
Qy 2661 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2720
Db 2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220
Qy 2721 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2780
Db 2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280

```
Qy       2781 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340

Qy       2841 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400

Qy       2901 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460

Qy       2961 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520

Qy       3021 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580

Qy       3081 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640

Qy       3141 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700

Qy       3201 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3261 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820

Qy       3321 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880

Qy       3381 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940

Qy       3441 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000

Qy       3501 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060

Qy       3561 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120

Qy       3621 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180

Qy       3681 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240

Qy       3741 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300

Qy       3801 TGCTGAAAAAAAA 3813
             |||||||||||||
Db       3301 TGCTGAAAAAAAA 3313
```

```
RESULT 25
AR361484
LOCUS       AR361484                4493 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 1 from patent US 6599728.
ACCESSION   AR361484
VERSION     AR361484.1  GI:33769332
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4493)
  AUTHORS   Morin,G.B., Funk,W.D. and Piatyszek,M.A.
  TITLE     Second mammalian tankyrase
  JOURNAL   Patent: US 6599728-A 1 29-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .4493
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           79.6%;  Score 3039;  DB 6;  Length 4493;
  Best Local Similarity   88.5%;  Pred. No. 0;
  Matches 3101;  Conservative    1;  Mismatches  375;  Indels   27;  Gaps    2;

Qy        337 AACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGA 396
              ||| | ||   |||||||||| ||||| |||| |||| || | |||||| ||||||||
Db        587 AACCATTCCNNAGGCTGGTGACCCCTGAAAAGGTNAACANCCNCAACACGGNGGGCAGGA 646

Qy        397 AATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGC 456
              |||||||||||||||||| |||||||||||||||||| || |  || |||||||||
Db        647 AATCCACCCCGCTGCACTTCCCCGCAGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGC 706

Qy        457 TTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATG 516
              |||| |||||||||||| ||||| ||| | || |||||||||||||||||||||||| ||
Db        707 TTCAAAATGGTGCAAATNTCCAANCACTTTATAATGGGGGCCTTATTCCTCTTCATANTG 766

Qy        517 CATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCA 576
              |||||||||||||||||| ||    ||||||||||||||||||||||||||||||||||
Db        767 CATGCTCTTTTGGTCATGCTAAAANTATCAATCTCCTTTTGCGACATGGTGCAGACCCCA 826

Qy        577 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 636
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 886

Qy        637 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 696
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 946

Qy        697 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 756
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 1006

Qy        757 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 816
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 1066

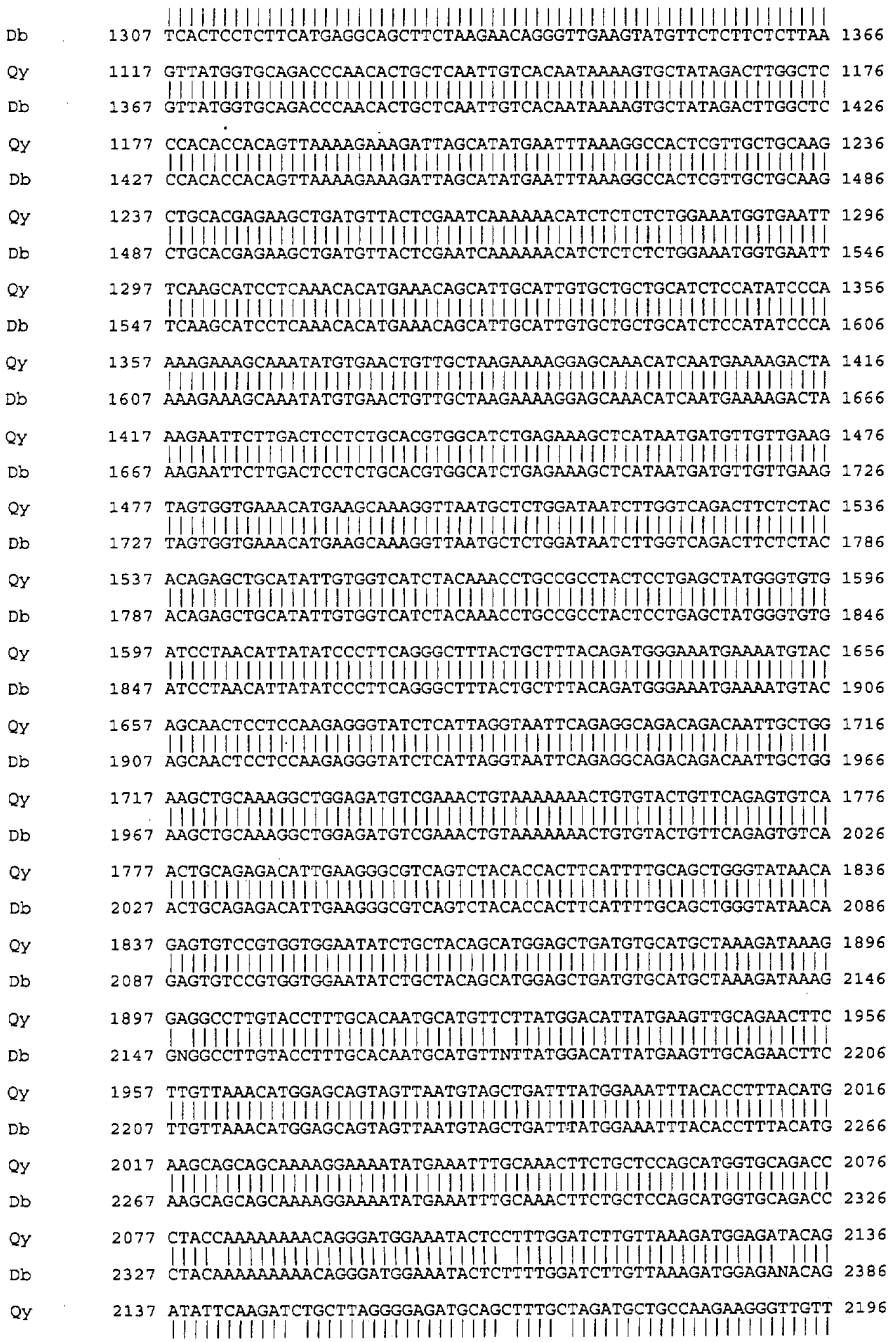
Qy        817 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 876
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 1126

Qy        877 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 936
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 1186

Qy        937 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 996
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 1246

Qy        997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1306

Qy       1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
```


Db 1307 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1366
Qy 1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
Db 1367 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1426
Qy 1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
Db 1427 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1486
Qy 1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
Db 1487 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1546
Qy 1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
Db 1547 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1606
Qy 1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
Db 1607 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1666
Qy 1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
Db 1667 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1726
Qy 1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
Db 1727 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1786
Qy 1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
Db 1787 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1846
Qy 1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
Db 1847 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1906
Qy 1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
Db 1907 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1966
Qy 1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
Db 1967 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 2026
Qy 1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836
Db 2027 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 2086
Qy 1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896
Db 2087 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 2146
Qy 1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956
Db 2147 GNGGCCTTGTACCTTTGCACAATGCATGTTNTTATGGACATTATGAAGTTGCAGAACTTC 2206
Qy 1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2016
Db 2207 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2266
Qy 2017 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2076
Db 2267 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2326
Qy 2077 CTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2136
Db 2327 CTACAAAAAAAAACAGGGATGGAAATACTCTTTTGGATCTTGTTAAAGATGGAGANACAG 2386
Qy 2137 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2196
Db 2387 ATATTCAAGATNTGCTTAGGGGAGATGCAGTTTTGTTAGATGCTGCCAAGAAGGGTTGTT 2446
Qy 2197 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2256
Db 2447 TAGCCAGAGTGAAGAAGTTNTNTTTTCCTGATAATGTAAATTGCCGNGATACCCAAGGCA 2506
Qy 2257 GACATTCAACACCTTTACATTTAGCAGCT------GGTTATAATAATTTAGAAGTTGCAG 2310
Db 2507 GACATTCAACACCTTTACATTTAGCAGGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2566
Qy 2311 AGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTT 2370
Db 2567 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2626
Qy 2371 TACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATG 2430
Db 2627 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2686
Qy 2431 CATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGG 2490
Db 2687 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2746
Qy 2491 GACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATC 2550
Db 2747 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2806
Qy 2551 AGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAG 2610
Db 2807 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAG 2866
Qy 2611 CAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGA 2670
Db 2867 CAGCCATGCCCCCATCTGTTCTGCCCTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGA 2926
Qy 2671 GAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTT 2730
Db 2927 GAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTT 2986
Qy 2731 CTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTA 2790
Db 2987 CTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTA 3046
Qy 2791 GTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATT 2850
Db 3047 GTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATT 3106
Qy 2851 TTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGA 2910
Db 3107 TTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTNAGA 3166
Qy 2911 GAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTG 2970
Db 3167 GAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTG 3226
Qy 2971 GAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCG 3030
Db 3227 GWATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCG 3286
Qy 3031 GACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTA 3090
Db 3287 GACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTA 3346
Qy 3091 TAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAG 3150
Db 3347 TAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAG 3406
Qy 3151 TTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCA 3210
Db 3407 TTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCA 3466
Qy 3211 AGATTCAGAAGGTTTGTAACAAGAAACTATG----------GGAAAGAT 3249
Db 3467 AGATTCAGAAGGTTTGTAACANNNNNNNNNNNGAGCCAAGATTCGGCACGAGGAAAGAT 3526

```
Qy       3250 ACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGC 3309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3527 ACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGC 3586

Qy       3310 TATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATG 3369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3587 TATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATG 3646

Qy       3370 CGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCA 3429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3647 CGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCA 3706

Qy       3430 ATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTT 3489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3707 ATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTT 3766

Qy       3490 GTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGC 3549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3767 GTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGC 3826

Qy       3550 AGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGC 3609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3827 AGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGC 3886

Qy       3610 CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATC 3669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3887 CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATC 3946

Qy       3670 CTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAG 3729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3947 CTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAG 4006

Qy       3730 TTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACG 3789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4007 TTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACG 4066

Qy       3790 TTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||
Db       4067 TTTTACTCCTTTGCTGAAAAAAAA 4090
```

RESULT 26
AX052530
LOCUS AX052530 5005 bp DNA linear PAT 12-JAN-2001
DEFINITION Sequence 9 from Patent WO0072021.
ACCESSION AX052530
VERSION AX052530.1 GI:12226767
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kuimov,A.B., Lagarkova,M.B., Koroleva,E.B., Turetskaya,R., Vdovichenko,K.B., Mescheryakov,A.O., Litchinitser,M.O., Kuprash,D., Nedospasov,S.B., Tureci,O., Sahin,U., Pfruendschuh,M., Oold,L.J., Knuth,A. and Jager,E.
TITLE Cancer associated antigens and uses therefor
JOURNAL Patent: WO 0072021-A 9 30-NOV-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)

FEATURES Location/Qualifiers
source 1. .5005
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
unsure 3838
/note="n = a, c, g or t"
unsure 3840
/note="n = a, c, g or t"
unsure 3842
/note="n = a, c, g or t"
unsure 3846
/note="n = a, c, g or t"
unsure 3850
/note="n = a, c, g or t"
unsure 3882
/note="n = a, c, g or t"
unsure 3916
/note="n = a, c, g or t"
unsure 3930
/note="n = a, c, g or t"
unsure 3931
/note="n = a, c, g or t"
unsure 3955
/note="n = a, c, g or t"

ORIGIN

Query Match 72.7%; Score 2774.6; DB 6; Length 5005;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2810; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

```
Qy        997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 60

Qy       1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 120

Qy       1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 180

Qy       1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240

Qy       1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 300

Qy       1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 360

Qy       1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
              |||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||||
Db        361 AAAGAAAGCAAATATGTGAACTGTTGCTAAAAAAAGGATCAAACATCAATGAAAAGACTA 420

Qy       1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480

Qy       1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540

Qy       1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 600

Qy       1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 660

Qy       1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 720

Qy       1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCA 780

Qy       1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 840
Qy       1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896
Db        841 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 900
Qy       1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956
Db        901 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 960
Qy       1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2016
Db        961 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1020
Qy       2017 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2076
Db       1021 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 1080
Qy       2077 CTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2136
Db       1081 CTACAAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 1140
Qy       2137 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2196
Db       1141 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 1200
Qy       2197 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2256
Db       1201 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 1260
Qy       2257 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 2316
Db       1261 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 1320
Qy       2317 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 2376
Db       1321 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 1380
Qy       2377 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 2436
Db       1381 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 1440
Qy       2437 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 2496
Db       1441 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 1500
Qy       2497 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 2556
Db       1501 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 1560
Qy       2557 GACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCA 2616
Db       1561 GACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCA 1620
Qy       2617 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 2676
Db       1621 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 1680
Qy       2677 CAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 2736
Db       1681 CAGGAGCCACTGCAGATGCTCTCT-TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 1739
Qy       2737 CCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 2796
Db       1740 CCAGCAGTCTTGACAAC-TATCTGGGAG-TTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 1797
Qy       2797 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 2856
Db       1798 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 1857
Qy       2857 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 2916
Db       1858 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 1917
```

```
Qy       2917 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 2976
Db       1918 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 1977
Qy       2977 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 3036
Db       1978 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 2037
Qy       3037 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 3096
Db       2038 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 2097
Qy       3097 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 3156
Db       2098 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 2157
Qy       3157 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 3216
Db       2158 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 2217
Qy       3217 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 3276
Db       2218 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 2277
Qy       3277 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 3336
Db       2278 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 2337
Qy       3337 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3396
Db       2338 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 2397
Qy       3397 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 3456
Db       2398 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 2457
Qy       3457 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 3516
Db       2458 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 2517
Qy       3517 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 3576
Db       2518 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 2577
Qy       3577 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 3636
Db       2578 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 2637
Qy       3637 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 3696
Db       2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 2697
Qy       3697 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 3756
Db       2698 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 2757
Qy       3757 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       2758 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2814
```

```
RESULT 27
AY142107
LOCUS       AY142107               3930 bp    mRNA    linear   VRT 30-MAY-2003
DEFINITION  Gallus gallus tankyrase 2 mRNA, complete cds.
ACCESSION   AY142107
VERSION     AY142107.1  GI:27461952
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
```

```
          Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 3930)
  AUTHORS   De Rycker,M., Venkatesan,R.N., Wei,C. and Price,C.M.
  TITLE     Vertebrate tankyrase domain structure and sterile alpha motif
            (SAM)-mediated multimerization
  JOURNAL   Biochem. J. 372 (Pt 1), 87-96 (2003)
  MEDLINE   22615216
   PUBMED   12589701
REFERENCE   2  (bases 1 to 3930)
  AUTHORS   De Rycker,M., Venkatesan,R.N. and Price,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2002) Department of Molecular Genetics,
            University of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH
            45267-0524, USA
FEATURES             Location/Qualifiers
     source          1. .3930
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
     CDS             171. .3674
                     /codon_start=1
                     /product="tankyrase 2"
                     /protein_id="AAN41650.1"
                     /db_xref="GI:27461953"
                     /translation="MAARRCAGGAAALAEAPGCGSAVEPARELFEACRNGDVERVKRL
                     VRPENVNSRDTAGRKSSPLHFAAGFGRKDVVEYLLQSGANVHARDDGGLIPLHNACSF
                     GHAEVVNLLLRHGANPNARDNWNYTPLHEAAIKGKTDVCIVLLQHGAEPTIRNTDGRT
                     ALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMSLLTPLNVNCHASDGRKSTPLHL
                     AAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDL
                     WQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSTIDLAPTPQLKERLAYEFKGH
                     SLLQAARESDVARIKKHLSLETVNFKHPQTHETALHCAAASPYPKRKQVCELLLRKGA
                     NINEKTKDFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAHCGHLQTC
                     RLLLSSGCDPSIVSLQGFTALQMGNESVQQLLQEGIPLGNSDADRQLLEAAKAGDVDT
                     VKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN
                     ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPPKKNR
                     DGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLCSPDNVNCRDTQGRHST
                     PLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVN
                     ATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVTADDVSALLTAA
                     MPPSALPSCYKPQVISVSQTASSTADPLSSVPSSPSSLSAASSLDNLSGSFSELPSVV
                     GTNSAEGATVLEKKEVSGVDFSINQFVRNLGLEHLIDIFEREQITLDVLVEMGHKELK
                     EIGINAYGHRHKIIKGVERLISGQQGLNPYLTLNTSSSGTLLIDLSSEDKEFQSVEEE
                     MQSTVREHRDGGHAGGVFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERML
                     FHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPIHKDR
                     SCYVCHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE
                     QAYPEYLITYQIVKPEATTEA"
ORIGIN

  Query Match           66.1%;  Score 2522.6;  DB 5;  Length 3930;
  Best Local Similarity  82.4%;  Pred. No. 0;
  Matches 2890;  Conservative    0;  Mismatches  619;  Indels    0;  Gaps    0;

Qy        217 CCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGG 276
              || | ||   | || | ||  ||   | ||||| ||  ||       | |    | |
Db        166 CCGGCATGGCGGCGCGGCGGTGCGCGGGCGGCGCGGCCGCCTTGGCTGAGGCGCCGGGCT 225

Qy        277 CCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGG 336
              ||   | | ||  |  | | ||||| ||||||||||||| ||| | ||||||||||||
Db        226 GCGGCTCGGCGGTGGAGCCGGCCCGGGAGCTGTTCGAGGCCTGCAGGAACGGGGACGTGG 285

Qy        337 AACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGA 396
              | || |||||| ||||||||  ||| |||||  ||||||||||||||||||||||||||||
Db        286 AGCGGGTCAAGCGGCTGGTGCGGCCCGAGAACGTGAACAGCCGCGACACGGCGGGCAGGA 345

Qy        397 AATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGC 456
              | |||| |||||||||||||||||||||||||||||||| ||||| || |||||  |||| || |
Db        346 AGTCCAGCCCGCTGCACTTCGCCGCAGGTTTTGGTCGGAAGGATGTAGTCGAATACTTAC 405

Qy        457 TTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATG 516
              | |||| |||||| |||||||| ||||| |||||||| ||| |||||||| || |||||  ||||
Db        406 TGCAGAGTGGTGCCAATGTCCACGCACGAGATGATGGAGGCCTTATCCCACTTCACAATG 465

Qy        517 CATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCA 576
              | |||||||||||||| |||||||| ||||||||||  |||| || |||||| | || |

Db        466 CTTGCTCTTTTGGTCACGCTGAAGTTGTCAATCTCCTCCTGCGGCACGGTGCAAATCCTA 525

Qy        577 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 636
              |||||| || |||||||||| ||||| || ||||| || || |||||  || ||  |  |
Db        526 ATGCTCGGGACAATTGGAATTACACTCCCCTTCATGAGGCAGCCATTAAGGGCAAAACAG 585

Qy        637 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 696
              ||||  |||||||  || | |||||||| ||||| |||||||||| || |||||||||
Db        586 ATGTCTGCATTGTATTGCTGCAGCATGGTGCTGAACCAACCATCCGGAACACAGATGGAA 645

Qy        697 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 756
              | ||||| |||||||||||||| |||||||| |||||||||| |||||||||| ||||
Db        646 GAACAGCTTTGGATTTAGCAGACCCATCTGCAAAAGCAGTGCTGACTGGTGAATACAAGA 705

Qy        757 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 816
              ||||||| |||||||||||||||||||||| |||||||||| |||||  |||| || |
Db        706 AAGATGAGCTCTTAGAAAGTGCCAGGAGTGGAAATGAAGAAAAGATGATGTCTCTTCTTA 765

Qy        817 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 876
              |||||||  |||||  ||||| || ||||||||||  |||||  |||||  | |
Db        766 CACCATTGAATGTTAACTGTCATGCGAGTGATGGCAGGAAGTCTACTCCTCTACATCTAG 825

Qy        877 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 936
              |||| || |||||| | |||||  || |||||||||||||||||||||||| ||||||| |
Db        826 CAGCTGGGTATAACCGTGTAAAAATAGTACAGCTGTTACTGCAACATGGGGCTGATGTGC 885

Qy        937 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 996
              | ||||| |||||  || ||| |||| ||| | ||||||||||||||| |||||| |||||||
Db        886 ACGCTAAGGATAAGGGAGATTTGGTGCCACTTCACAATGCCTGTTCCTATGGCCATTATG 945

Qy        997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
              ||||||| || ||| | || |||||||| || ||||| ||||| |||||  |||||||  |
Db        946 AAGTAACAGAGCTTCTAGTGAAGCATGGAGCATGTGTGAATGCGATGGATCTGTGGCAGT 1005

Qy       1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
              |||||||||| ||||| |||||||| |||||||||| |||||||| |||||| | ||||
Db       1006 TCACTCCTCTGCATGAAGCAGCTTCAAAGAACAGGGTGGAAGTATGCTCTCTTTTGTTAA 1065

Qy       1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
              |||||||||| |||||||||| | || |||||||| |||||  | || || |||||||
Db       1066 GTTATGGTGCTGACCCAACACTGTTGAACTGTCACAACAAAAGCACCATTGATTTGGCTC 1125

Qy       1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
              | || || || |||||||| |||||||||||||||| ||||| |||||  |||| ||||
Db       1126 CAACGCCTCAATTAAAAGAACGATTAGCATATGAATTCAAAGGTCACTCACTGCTACAAG 1185

Qy       1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
              ||||| ||||| | |||||| |||| || || || ||||| ||| ||||||  || || |
Db       1186 CTGCAAGAGAATCAGATGTTGCTCGTATAAAGAAGCATCTTTCTTTGTGGAAACAGTTAACT 1245

Qy       1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
              |||||||||||||||| ||||| ||||| |||| |||||||||||||||||| ||||| |
Db       1246 TCAAGCATCCTCAAACGCATGAGACAGCACTGCACTGTGCTGCTGCATCTCCTTATCCTA 1305

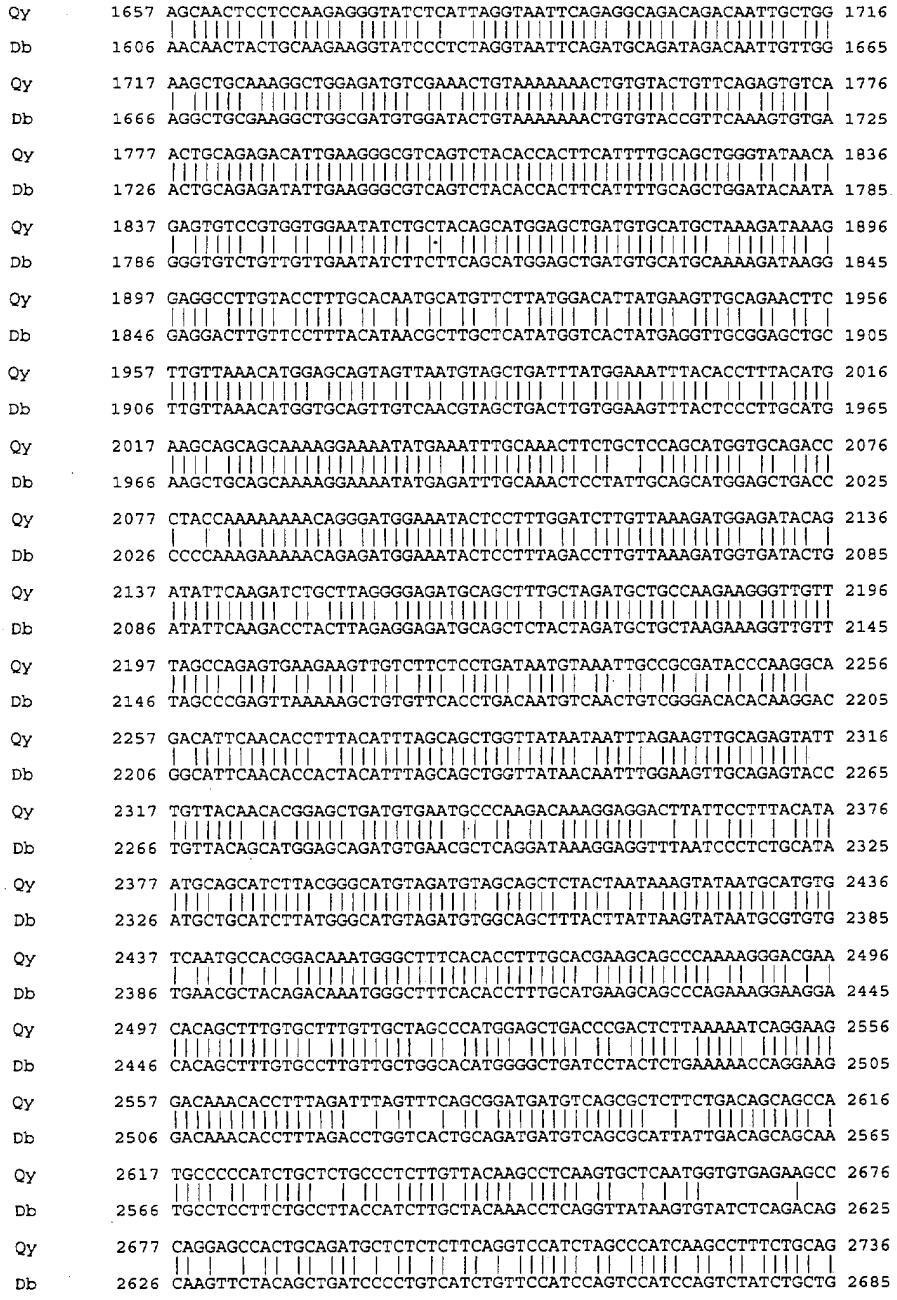
Qy       1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
              | ||||| ||| |||||||| || |||| || |||||||| ||||| |||||||||| |
Db       1306 AGAGAAAACAAGTATGTGAATTGCTGCTGAGGAAAGGAGCTAACATTAATGAAAAGACAA 1365

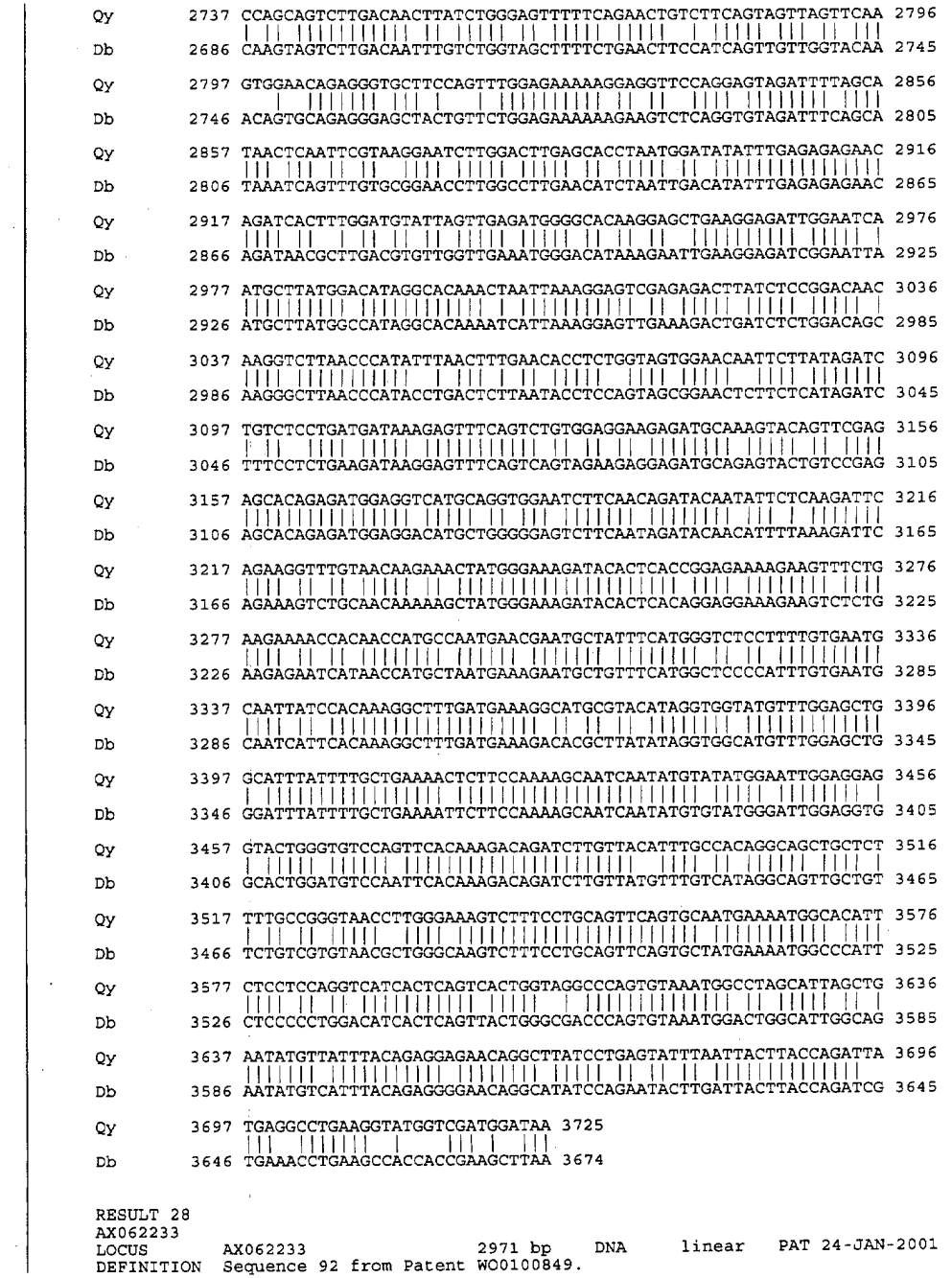
Qy       1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
              |||| |||||||||||||||| |||||||| || |||||||||||||||||||||||||
Db       1366 AAGACTTCTTGACTCCTCTGCATGTGGCATCAGAAAAAGCTCATAATGATGTTGTTGAAG 1425

Qy       1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
              |||| |||||||||||||||||||||| ||||||| |||| ||||| ||||| |
Db       1426 TAGTTGTGAAACATGAAGCAAAGGTTAATGCATTGGATAACCTTGGCCAGACCTCTCTTC 1485

Qy       1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
              ||||||| ||| |||||||||| || || || ||| | |  | |||| | ||| ||||
Db       1486 ACAGAGCGGCACATTGTGGTCACCTCCAGACTTGCAGGTTGTTGTTGAGTTCTGGATGTG 1545

Qy       1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
              |||||  |||| | ||  | |||||||||| || |||| ||||| ||||||| ||| |
Db       1546 ATCCTTCCATTGTGTCTTTGCAGGGCTTTACAGCCTTACAAATGGGGAATGAAAGTGTGC 1605
```





RESULT 28
AX062233
LOCUS AX062233 2971 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 92 from Patent WO0100849.

```
ACCESSION   AX062233
VERSION     AX062233.1  GI:12540134
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 92 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2971
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

Query Match 63.9%; Score 2439.8; DB 6; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1371 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1430
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1431 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1490
Db            61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1491 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
Db           121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1551 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1610
Db           181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1670
Db           241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1671 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1730
Db           301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1731 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1790
Db           361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1791 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1850
Db           421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1851 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1910
Db           481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1911 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1970
Db           541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1971 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2030
Db           601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2031 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAAC 2090
Db           661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAAC 720

Qy       2091 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2150
Db           721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2151 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2210
Db           781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2211 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2270
Db           841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2271 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2330
Db           901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2331 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2390
Db           961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2450
Db          1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2451 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2510
Db          1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy       2511 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2570
Db          1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy       2571 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630
Db          1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2631 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2690
Db          1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy       2691 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2750
Db          1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy       2751 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2810
Db          1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy       2811 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2870
Db          1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy       2871 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930
Db          1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy       2931 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2990
Db          1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2991 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3050
Db          1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy       3051 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3110
Db          1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy       3111 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3170
Db          1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy       3171 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3230
Db          1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860
```

```
Qy       3231 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy       3291 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy       3351 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy       3411 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy       3471 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3530
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy       3531 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3590
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3591 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3650
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3711 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3770
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy       3771 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 29
AX062229
LOCUS AX062229 3353 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 88 from Patent WO0100849.
ACCESSION AX062229
VERSION AX062229.1 GI:12540130
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 88 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .3353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS <1. .2355
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC25057.1"
/db_xref="GI:12540131"
/db_xref="REMTREMBL:CAC25057"
/translation="CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV
NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEG
ISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVV
EYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA
AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCL
ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPL
HNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK
NQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSP
SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL
MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNT
SGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW
ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAEN
SSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGH
HSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG"
ORIGIN

Query Match 63.9%; Score 2439.8; DB 6; Length 3353;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1371 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1430
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1431 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1490
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1491 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1551 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1670
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1671 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1730
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1731 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1790
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1791 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1850
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1851 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1910
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1911 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1970
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1971 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2030
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2031 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2090
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAAC 720

Qy       2091 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2151 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2210
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2211 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2270
```

```
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900
Qy       2271 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2330
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960
Qy       2331 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2390
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020
Qy       2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2450
Db       1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080
Qy       2451 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2510
Db       1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140
Qy       2511 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2570
Db       1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200
Qy       2571 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630
Db       1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260
Qy       2631 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2690
Db       1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320
Qy       2691 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2750
Db       1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380
Qy       2751 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2810
Db       1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440
Qy       2811 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2870
Db       1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500
Qy       2871 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930
Db       1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560
Qy       2931 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2990
Db       1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620
Qy       2991 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3050
Db       1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680
Qy       3051 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3110
Db       1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740
Qy       3111 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3170
Db       1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800
Qy       3171 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3230
Db       1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860
Qy       3231 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3290
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920
Qy       3291 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3350
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980
Qy       3351 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3410
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040
Qy       3411 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3470
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100
Qy       3471 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3530
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160
Qy       3531 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3590
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220
Qy       3591 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3650
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280
Qy       3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340
Qy       3711 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3770
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400
Qy       3771 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

```
RESULT 30
AX062231
LOCUS       AX062231                3799 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 90 from Patent WO0100849.
ACCESSION   AX062231
VERSION     AX062231.1  GI:12540132
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 90 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3799
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <3. .2273
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25058.1"
                     /db_xref="GI:12540133"
                     /db_xref="REMTREMBL:CAC25058"
                     /translation="AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS
                     YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC
                     TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG
                     HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTP
                     LDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA
                     AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW
                     AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA
                     LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGT
                     EGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN
                     AYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV
                     REHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSP
```

FVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYIC
HRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPE
YLITYQIMRPEGMVDG"

ORIGIN

Query Match 61.8%; Score 2357.8; DB 6; Length 3799;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1453 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 1512
Db             1 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 60

Qy       1513 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 1572
Db            61 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 120

Qy       1573 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 1632
Db            121 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 180

Qy       1633 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 1692
Db            181 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 240

Qy       1693 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1752
Db            241 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 300

Qy       1753 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 1812
Db            301 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 360

Qy       1813 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 1872
Db            361 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 420

Qy       1873 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 1932
Db            421 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 480

Qy       1933 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 1992
Db            481 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 540

Qy       1993 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 2052
Db            541 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 600

Qy       2053 TTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGG 2112
Db            601 TTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGG 660

Qy       2113 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 2172
Db            661 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 720

Qy       2173 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 2232
Db            721 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 780

Qy       2233 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 2292
Db            781 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 840

Qy       2293 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2352
Db            841 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 900

Qy       2353 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2412
Db            901 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 960

Qy       2413 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2472
Db             961 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 1020

Qy       2473 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2532
Db            1021 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 1080

Qy       2533 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2592
Db            1081 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 1140

Qy       2593 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2652
Db            1141 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 1200

Qy       2653 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2712
Db            1201 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 1260

Qy       2713 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2772
Db            1261 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 1320

Qy       2773 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2832
Db            1321 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 1380

Qy       2833 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2892
Db            1381 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 1440

Qy       2893 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2952
Db            1441 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 1500

Qy       2953 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 3012
Db            1501 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 1560

Qy       3013 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 3072
Db            1561 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 1620

Qy       3073 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 3132
Db            1621 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 1680

Qy       3133 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 3192
Db            1681 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 1740

Qy       3193 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 3252
Db            1741 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 1800

Qy       3253 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 3312
Db            1801 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 1860

Qy       3313 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 3372
Db            1861 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 1920

Qy       3373 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 3432
Db            1921 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 1980

Qy       3433 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 3492
Db            1981 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 2040

Qy       3493 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 3552
```

```
Db       2041 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 2100

Qy       3553 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 3612

Db       2101 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 2160

Qy       3613 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 3672

Db       2161 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 2220

Qy       3673 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 3732

Db       2221 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 2280

Qy       3733 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 3792

Db       2281 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 2340

Qy       3793 TACTCCTTTGCTGAAAAAAAA 3813

Db       2341 TACTCCTTTGCTGAAAAAAAA 2361


RESULT 31
AX057580
LOCUS       AX057580                3254 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 3 from Patent WO0077225.
ACCESSION   AX057580
VERSION     AX057580.1  GI:12310302
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Chi,N.W. and Lodish,H.F.
  TITLE     A novel insulin signaling molecule
  JOURNAL   Patent: WO 0077225-A 3 21-DEC-2000;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; THE GENERAL
            HOSPITAL CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3254
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match           55.6%;  Score 2122;  DB 6;  Length 3254;
  Best Local Similarity   77.7%;  Pred. No. 0;
  Matches 2895;  Conservative    0;  Mismatches  275;  Indels  558;  Gaps    5;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Db         82 CGCGCTGCCTCCACCGCGGTGGGGCAGCCGGGGCTCACGGAGCCGAGCCGGGGGCGCGCG 141

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Db        142 GCGACGCGG-CCATGGGACCGTGCCGACTCTGGTGACAGCGGGGAGCGGAGCGGCCCGGG 200

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180

Db        201 CCCTGAGCGCGTCTCCTGCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCCGGGCTCCTGC 260

Qy        181 TCCGGTTGCTGGCGCTGTTGCT---GGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCC 237

Db        261 TGCCTTTGCTGGCGCTGCTGCTGGCGGCGGCGGCGGCGGCCCGGATCATGTCTGGCCGCC 320

Qy        238 GCTGCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCG 297

Db        321 GCTGCGCCGGCGGGGGCGCGGCCTGCGCGAGCGCCGGGGCCGAGGCCGTGGAGCCGTCCG 380

Qy        298 CCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGA 357

Db        381 CCCGCGAGCTGTTCGAGGCTTGCCGCAACGGGGACGTGGAGCGAGTCAAGAGGCTGGTGA 440

Qy        358 CGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCG 417

Db        441 CGCCCGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAGTCCACCCCGCTCCACTTCG 500

Qy        418 CCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCC 477

Db        501 CCGCAGGTTTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGC 560

Qy        478 AAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTG 537

Db        561 AAGCGCGTGATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTG 620

Qy        538 AAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATT 597

Db        621 AAGTAGTCAATCTTCTTTTGCAACACGGTGCAGACCCAAATGCTCGGGATAATTGGAATT 680

Qy        598 ATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTAC 657

Db        681 ATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTAC 740

Qy        658 AGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAG 717

Db        741 AGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAG 800

Qy        718 ATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTG 777

Db        801 ACCCATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGAAAGATGAACTTCTGGAAAGTG 860

Qy        778 CCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCC 837

Db        861 CCAGGAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCC 920

Qy        838 ACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAA 897

Db        921 ATGCAAGTGATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGATACAACAGAGTAA 980

Qy        898 AGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATC 957

Db        981 AAATTGTACAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATC 1040

Qy        958 TGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCA 1017

Db       1041 TGGTACCACTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCA 1100

Qy       1018 AGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAG 1077

Db       1101 AGCACGGTGCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAG 1160

Qy       1078 CTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACAC 1137

Db       1161 CTTCTAAGAACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCCAACCC 1220

Qy       1138 TGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAA 1197

Db       1221 TGCTGAACTGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAA 1280

Qy       1198 GATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTA 1257

Db       1281 GATTATCGTATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTA 1340

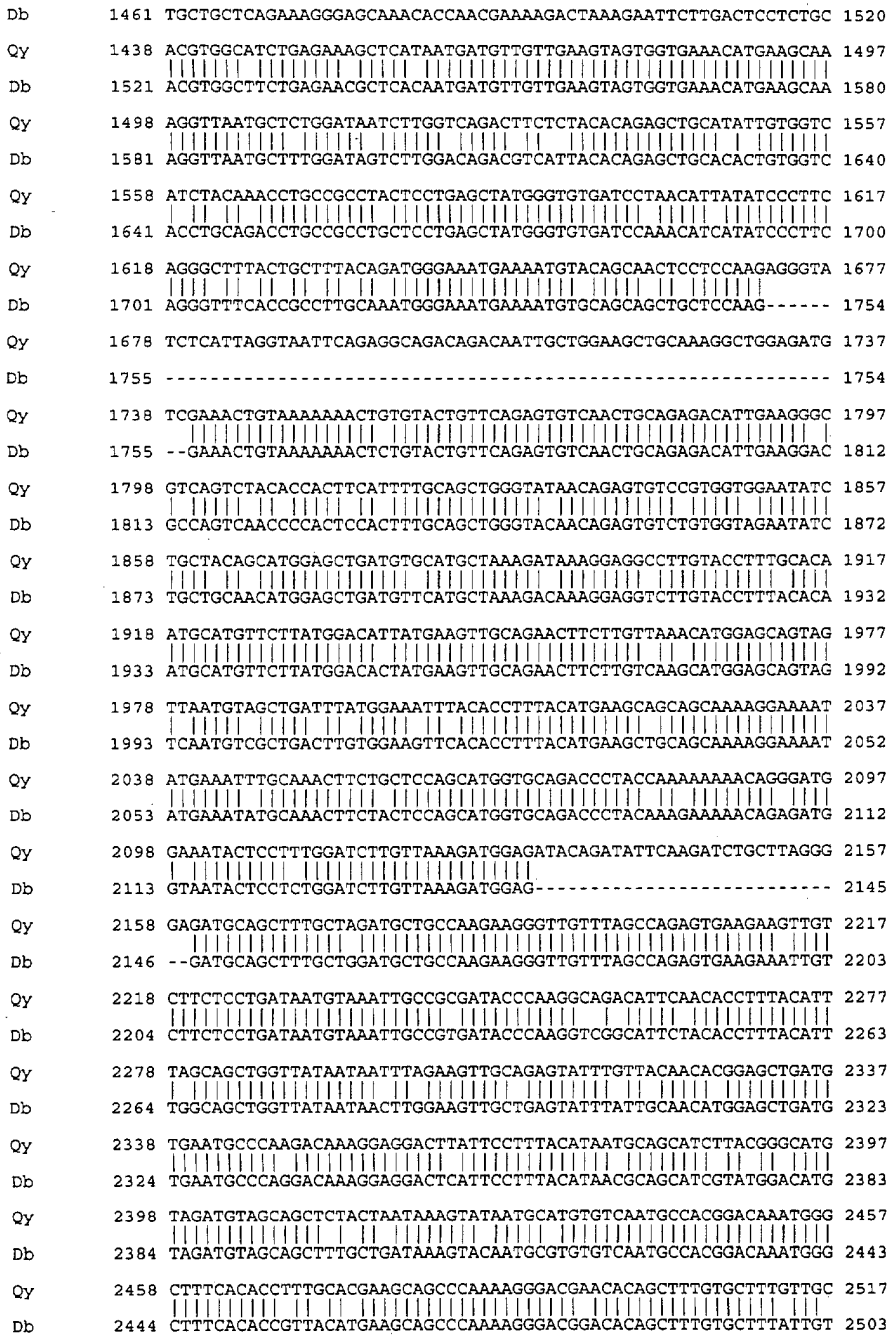
Qy       1258 CTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATG 1317

Db       1341 CAAGAATCAAAAAACATCTTTCTCTGGAAATGGTGAATTTCAAACATCCTCAGACACATG 1400

Qy       1318 AAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAAC 1377

Db       1401 AAACAGCACTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAGATATGTGAAC 1460

Qy       1378 TGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGC 1437
```


Db 1461 TGCTGCTCAGAAAGGGAGCAAACACCAACGAAAAGACTAAAGAATTCTTGACTCCTCTGC 1520
Qy 1438 ACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAA 1497
Db 1521 ACGTGGCTTCTGAGAACGCTCACAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAA 1580
Qy 1498 AGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTC 1557
Db 1581 AGGTTAATGCTTTGGATAGTCTTGGACAGACGTCATTACACAGAGCTGCACACTGTGGTC 1640
Qy 1558 ATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTC 1617
Db 1641 ACCTGCAGACCTGCCGCCTGCTCCTGAGCTATGGGTGTGATCCAAACATCATATCCCTTC 1700
Qy 1618 AGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTA 1677
Db 1701 AGGGTTTCACCGCCTTGCAAATGGGAAATGAAAATGTGCAGCAGCTGCTCCAAG------ 1754
Qy 1678 TCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATG 1737
Db 1755 ---------- 1754
Qy 1738 TCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGC 1797
Db 1755 --GAAACTGTAAAAAAACTCTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGAC 1812
Qy 1798 GTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATC 1857
Db 1813 GCCAGTCAACCCCACTCCACTTTGCAGCTGGGTACAACAGAGTGTCTGTGGTAGAATATC 1872
Qy 1858 TGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACA 1917
Db 1873 TGCTGCAACATGGAGCTGATGTTCATGCTAAAGACAAAGGAGGTCTTGTACCTTTACACA 1932
Qy 1918 ATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAG 1977
Db 1933 ATGCATGTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAG 1992
Qy 1978 TTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAAT 2037
Db 1993 TCAATGTCGCTGACTTGTGGAAGTTCACACCTTTACATGAAGCTGCAGCAAAAGGAAAAT 2052
Qy 2038 ATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATG 2097
Db 2053 ATGAAATATGCAAACTTCTACTCCAGCATGGTGCAGACCCTACAAAGAAAAACAGAGATG 2112
Qy 2098 GAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGG 2157
Db 2113 GTAATACTCCTCTGGATCTTGTTAAAGATGGAG---------- 2145
Qy 2158 GAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGT 2217
Db 2146 --GATGCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAATTGT 2203
Qy 2218 CTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATT 2277
Db 2204 CTTCTCCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACACCTTTACATT 2263
Qy 2278 TAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATG 2337
Db 2264 TGGCAGCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGGAGCTGATG 2323
Qy 2338 TGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATG 2397
Db 2324 TGAATGCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATG 2383
Qy 2398 TAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGG 2457
Db 2384 TAGATGTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGG 2443
Qy 2458 CTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGC 2517
Db 2444 CTTTCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGT 2503
Qy 2518 TAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAG 2577
Db 2504 TGGCCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAG 2563
Qy 2578 TTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCT 2637
Db 2564 TTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCC---------- 2606
Qy 2638 CTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTC 2697
Db 2607 ---------- 2606
Qy 2698 TCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTAT 2757
Db 2607 ---------- 2606
Qy 2758 CTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCA 2817
Db 2607 ---------- 2606
Qy 2818 GTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATC 2877
Db 2607 ---------- 2606
Qy 2878 TTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAG 2937
Db 2607 ---------- 2606
Qy 2938 TTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACA 2997
Db 2607 ---------- 2606
Qy 2998 AACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAA 3057
Db 2607 ---------- 2606
Qy 3058 CTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGT 3117
Db 2607 ----------GGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAAT 2646
Qy 3118 TTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATG 3177
Db 2647 TTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACG 2706
Qy 3178 CAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAAC 3237
Db 2707 CAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAAT 2766
Qy 3238 TATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCA 3297
Db 2767 TGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCAA 2826
Qy 3298 ATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTG 3357
Db 2827 ACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTG 2886
Qy 3358 ATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACT 3417
Db 2887 ATGAAAGGCATGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACT 2946
Qy 3418 CTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACA 3477
Db 2947 CTTCCAAAAGCAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTCCAATTCACA 3006
Qy 3478 AAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAA 3537
Db 3007 AAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAA 3066
Qy 3538 AGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAG 3597
Db 3067 AGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGG 3126

```
Qy       3598 TCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAG 3657
             |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3127 TCACTGGCCGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAG 3186

Qy       3658 AACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCG 3717
             |||||||||||||||| ||| |||||||||||||||| | |||||||||||||||||| |
Db       3187 AACAGGCTTATCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAGGTATGGTTG 3246

Qy       3718 ATGGATAA 3725
             |||| | |
Db       3247 ATGGCTGA 3254


RESULT 32
HSM804989
LOCUS       HSM804989              4130 bp    mRNA    linear   PRI 10-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp667A1222 (from clone DKFZp667A1222).
ACCESSION   AL833676
VERSION     AL833676.1  GI:21734324
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4130)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp667A1222) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .4130
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp667A1222"
                     /db_xref="taxon:9606"
                     /clone="DKFZp667A1222"
                     /tissue_type="lymph node"
                     /clone_lib="667 (synonym: hlno2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     polyA_signal    4094. .4099
     polyA_site      4113
ORIGIN

  Query Match           50.4%;  Score 1922.4;  DB 9;  Length 4130;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1923;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1890 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1949
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 60

Qy       1950 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 2009
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 120

Qy       2010 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2069
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 180

Qy       2070 GCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 2129
             |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 240

Qy       2130 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 2189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 300

Qy       2190 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2249
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 360

Qy       2250 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2309
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 420

Qy       2310 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2369
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 480

Qy       2370 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2429
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 540

Qy       2430 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 600

Qy       2490 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 660

Qy       2550 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 720

Qy       2610 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 780

Qy       2670 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 840

Qy       2730 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 900

Qy       2790 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 960

Qy       2850 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 1020

Qy       2910 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 1080

Qy       2970 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 3029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 1140

Qy       3030 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 3089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 1200

Qy       3090 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 3149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 1260

Qy       3150 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 1320

Qy       3210 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1321 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 1380

Qy      3270 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3329
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 1440

Qy      3330 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 1500

Qy      3390 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 1560

Qy      3450 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 1620

Qy      3510 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 1680

Qy      3570 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 1740

Qy      3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 1800

Qy      3690 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 1860

Qy      3750 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3809
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 1920

Qy      3810 AAAA 3813
             ||||
Db      1921 AAAA 1924
```

RESULT 33
BC057370

```
LOCUS       BC057370                4054 bp    mRNA    linear   ROD 08-OCT-2003
DEFINITION  Mus musculus tankyrase, TRF1-interacting ankyrin-related ADP-ribose
            polymerase, mRNA (cDNA clone MGC:67299 IMAGE:6835784), complete
            cds.
ACCESSION   BC057370
VERSION     BC057370.1  GI:34980998
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4054)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 4054)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 126 Row: p Column: 6
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.
FEATURES             Location/Qualifiers
     source          1. .4054
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="MGC:67299 IMAGE:6835784"
                     /tissue_type="Brain, enriched mouse brain 12.5dpc"
                     /clone_lib="NIH_BMAP_FI0"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
     gene            1. .4054
                     /gene="Tnks"
                     /db_xref="LocusID:21951"
                     /db_xref="MGI:1341087"
     CDS             7. .3969
                     /codon_start=1
                     /product="Tnks protein"
                     /protein_id="AAH57370.1"
                     /db_xref="GI:34980999"
                     /db_xref="LocusID:21951"
                     /translation="MAASRRSQHHHHHQQQLQPAPGASAPPPPPPPLSPGLAPGPT
                     PASPTAGGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGAVCTVAAPAAVPAASA
                     AVGVAPTPAGGGGGGNNSASSASSPTSSSSSSPSSPGSSLAESPEAAGVGSTATLGA
                     GAAGLGPGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG
                     FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNY
                     TPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLE
                     AARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKD
                     KGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSH
                     GADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAARE ADLAKVKKTLALEIIN
                     FKQPQSHETALHCAVASLHPKRKQVAELLLRKGANVNEKNKDFMTPLHVAAERAHNDV
                     MEVLHKHGAKMNALDSLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMG
                     NEAVQQILSESTPMRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHSTPLH
                     FAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVAD
                     LWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLRGDA
```

```
                     ALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADV
                     NAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCAL
                     LLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVVSASLISP
                     ASTPSCLSAASSIDNLTGPLTDLAVGGASNAGDGAAGAERKEGEVAGLDMNISQFLKS
                     LGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNP
                     YLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKV
                     VNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGMFGAG
                     IYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFSTMKMA
                     HAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTATAAEQKT
                     "
     misc_feature    538. .1032
                     /note="Arp; Region: FOG: Ankyrin repeat [General function
                     prediction only]"
                     /db_xref="CDD:COG0666"
     misc_feature    3073. .3246
                     /note="SAM; Region: Sterile alpha motif. Widespread domain
                     in signalling and nuclear proteins. In EPH-related
                     tyrosine kinases, appears to mediate cell-cell initiated
                     signal transduction via the binding of SH2-containing
                     proteins to a conserved tyrosine that is phosphorylated.
                     In many cases mediates homodimerisation"
                     /db_xref="CDD:smart00454"
     misc_feature    3424. .3903
                     /note="PARP; Region: Poly(ADP-ribose) polymerase catalytic
                     domain. Poly(ADP-ribose) polymerase catalyses the covalent
                     attachment of ADP-ribose units from NAD+ to itself and to
                     a limited number of other DNA binding proteins, which
                     decreases their affinity for DNA. Poly(ADP-ribose)
                     polymerase is a regulatory component induced by DNA
                     damage. The carboxyl-terminal region is the most highly
                     conserved region of the protein. Experiments have shown
                     that a carboxyl 40 kDa fragment is still catalytically
                     active"
                     /db_xref="CDD:pfam00644"
ORIGIN

  Query Match           50.4%;  Score 1921.4;  DB 10;  Length 4054;
  Best Local Similarity  72.7%;  Pred. No. 0;
  Matches 2521;  Conservative    0;  Mismatches  921;  Indels   27;  Gaps    2;

Qy      246 GGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAG 305
Db      481 GGAGCCGGGGCGGCGGGCCTCGGCCCGGGGGTCCCAGCCGTGAGTGGGGCCCTGCGGGAG 540

Qy      306 CTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAG 365
Db      541 CTGCTCGAGGCCTGTCGGAATGGGGACGTGTCCCGGGTGAAGAGGTTGGTGGACGCCGCG 600

Qy      366 AAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGT 425
Db      601 AACGTGAATGCCAAGGACATGGCCGGCCGAAAGTCTTCTCCCCTGCACTTCGCAGCAGGT 660

Qy      426 TTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGT 485
Db      661 TTTGGAAGGAAGGATGTTGTAGAACACTTGCTGCAGATGGGTGCTAATGTCCACGCCCGT 720

Qy      486 GATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTC 545
Db      721 GATGACGGGGGTCTCATCCCACTTCATAATGCCTGCTCTTTCGGCCATGCTGAGGTTGTG 780

Qy      546 AATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCT 605
Db      781 AGTCTGTTACTGTGTCAAGGAGCAGATCCGAATGCCAGAGATAACTGGAACTATACACCA 840

Qy      606 CTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGA 665
Db      841 CTGCATGAAGCTGCCATTAAAGGGAAAATTGATGTCTGTATTGTGCTGCTGCAGCACGGA 900

Qy      666 GCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCT 725
Db      901 GCCGATCCAAACATCCGGAACACCGATGGGAAATCAGCTCTGGACCTGGCAGATCCTTCA 960

Qy      726 GCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGT 785

Db      961 GCAAAAGCTGTCCTCACAGGTGAATACAAAAAAGACGAACTCCTAGAAGCTGCTAGGAGT 1020

Qy      786 GGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGT 845
Db     1021 GGTAATGAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGT 1080

Qy      846 GATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTA 905
Db     1081 GATGGACGAAAGTCTACTCCGTTACACCTGGCAGCAGGCTACAACAGAGTTCGGATAGTT 1140

Qy      906 CAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCA 965
Db     1141 CAGCTGCTTCTACAGCATGGTGCTGATGTCCACGCGAAAGACAAAGGTGGACTTGTACCT 1200

Qy      966 TTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGT 1025
Db     1201 CTTCATAATGCATGCTCATACGGACACTATGAGGTCACAGAACTGCTGCTAAAGCATGGA 1260

Qy     1026 GCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAG 1085
Db     1261 GCTTGTGTCAACGCTATGGATCTGTGGCAGTTCACACCGCTGCATGAAGCTGCTTCAAAG 1320

Qy     1086 AACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAAT 1145
Db     1321 AACCGGGTAGAAGTCTGTTCTCTGTTACTTAGTCACGGTGCCGATCCAACCTTGGTCAAT 1380

Qy     1146 TGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCA 1205
Db     1381 TGCCATGGCAAAAGTGCTGTCGACATGGCTCCAACTCCTGAGCTCCGGGAAAGACTGACA 1440

Qy     1206 TATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATC 1265
Db     1441 TATGAATTTAAAGGCCATTCTTTACTCCAAGCCGCCAGAGAAGCAGACCTGGCCAAAGTT 1500

Qy     1266 AAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCA 1325
Db     1501 AAAAAAACGCTTGCTTTGGAAATCATTAACTTCAAGCAGCCACAGTCTCATGAAACTGCA 1560

Qy     1326 TTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTA 1385
Db     1561 TTGCACTGTGCTGTGGCCTCCTTGCATCCCAAACGAAAGCAGGTGGCAGAGCTGTTACTG 1620

Qy     1386 AGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCA 1445
Db     1621 AGGAAGGGAGCCAACGTCAACGAAAAGAACAAGGATTTCATGACCCCTCTGCATGTTGCA 1680

Qy     1446 TCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAAT 1505
Db     1681 GCTGAAAGAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCACGGAGCAAAGATGAAT 1740

Qy     1506 GCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAA 1565
Db     1741 GCACTGGATAGTCTTGGTCAGACTGCTTTGCACAGAGCTGCCTTAGCGGGCCACCTGCAG 1800

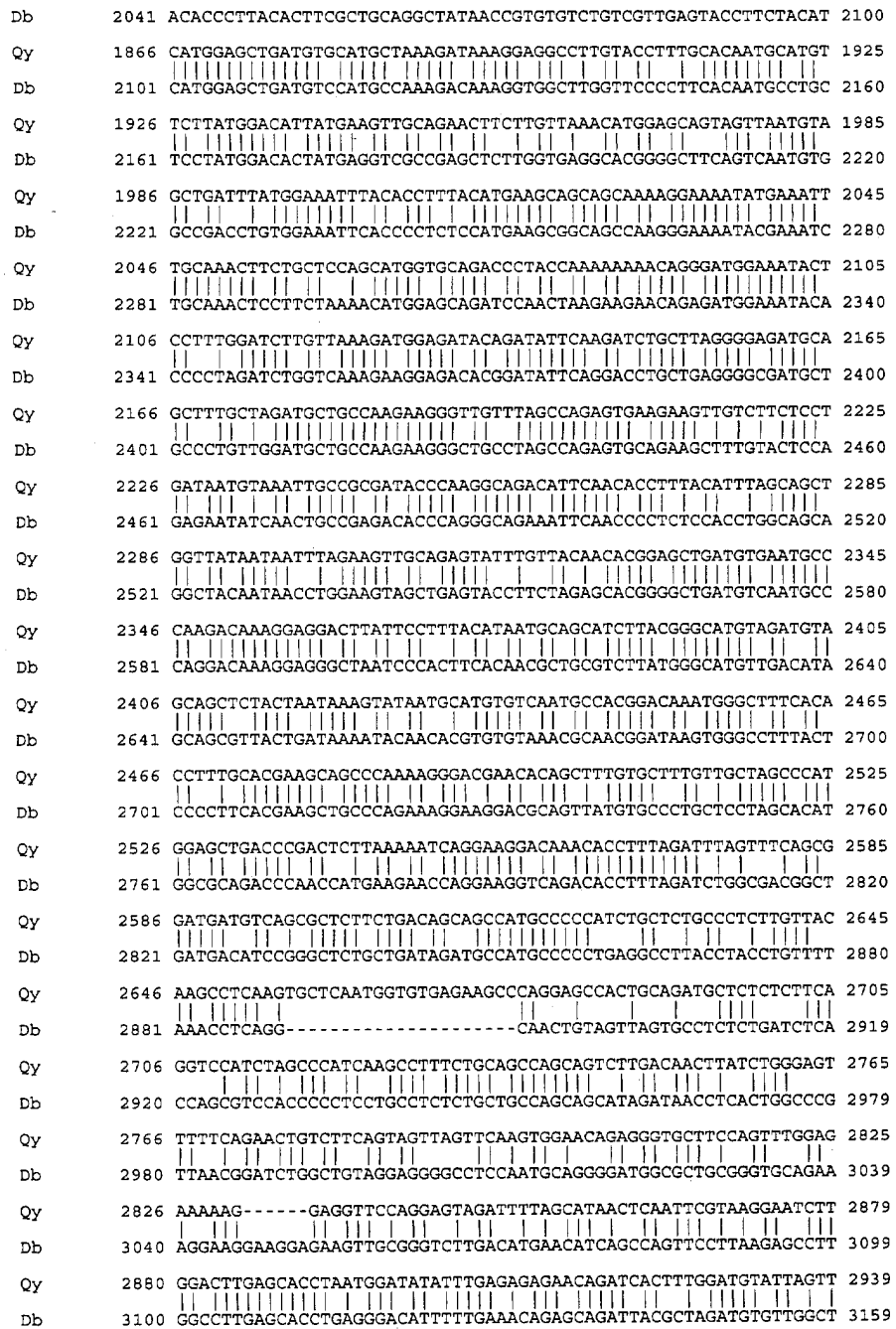
Qy     1566 ACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTT 1625
Db     1801 ACCTGCCGCCTCCTGCTGAGTTATGGCTCTGATCCTTCCATCATCTCCCTGCAAGGCTTC 1860

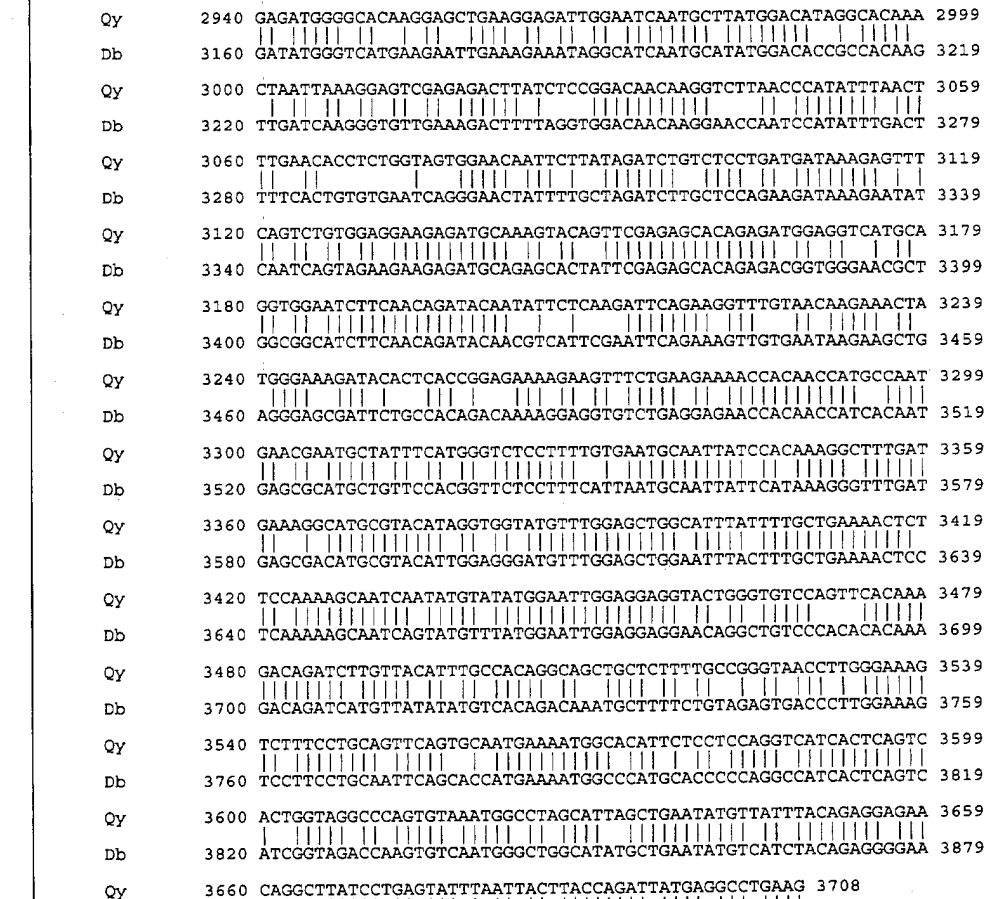
Qy     1626 ACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTA 1685
Db     1861 ACGGCAGCGCAGATGGGGAATGAAGCAGTGCAACAGATTCTGAGCGAGAGTACTCCTATG 1920

Qy     1686 GGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACT 1745
Db     1921 CGCACTTCGGACGTGGACTATCGTCTCTTAGAGGCATCGAAAGCTGGTGACTTGGAAACT 1980

Qy     1746 GTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCT 1805
Db     1981 GTGAAGCAACTCTGCAGCCCGCAGAATGTGAACTGCAGAGACCTGGAGGGCCGGCACTCT 2040

Qy     1806 ACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAG 1865
```





RESULT 34
AY142108
LOCUS AY142108 3804 bp mRNA linear VRT 30-MAY-2003
DEFINITION Gallus gallus tankyrase 1 mRNA, complete cds.
ACCESSION AY142108
VERSION AY142108.1 GI:27461954
KEYWORDS .
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 3804)
AUTHORS De Rycker,M., Venkatesan,R.N., Wei,C. and Price,C.M.
TITLE Vertebrate tankyrase domain structure and sterile alpha motif (SAM)-mediated multimerization
JOURNAL Biochem. J. 372 (Pt 1), 87-96 (2003)
MEDLINE 22615216
PUBMED 12589701

```
REFERENCE   2  (bases 1 to 3804)
  AUTHORS   Venkatesan,R.N., De Rycker,M. and Price,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2002) Department of Molecular Genetics,
            University of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH
            45267-0524, USA
FEATURES             Location/Qualifiers
     source          1. .3804
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
     CDS             1. .3801
                     /codon_start=1
                     /product="tankyrase 1"
                     /protein_id="AAN41651.1"
                     /db_xref="GI:27461955"
                     /translation="MAAPPRRSQHHHHHHGPPPPPGPASPPAAASPPRSPSLAPAELG
                     PAAQRHSLAGPEGEAPPDAERPPAPECSEGAAPGPPPGSGSSSSGSSASSSSSSSSTSS
                     SVASSPAAESPEAAGPSGAFRELLEACRNGDVTRVKRLVDAGNVNAKDMAGRKSTPLH
                     FAAGFGRKDVVEHLLQTGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARD
                     NWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAEAVLTGEYKKD
                     ELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADV
                     HAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSL
                     LLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLAL
                     EIINFKQPQSHETALHCAVAAVHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAEKA
                     HNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLNYGSDPSIISLQGFTA
                     AQIGNEAVQQILSESTPVRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHS
                     TPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASV
                     NVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLL
                     RGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEH
                     GADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQ
                     LCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKLQATVVSAS
                     LISPASTPSCLSAASSIDNLTGPLAELAVGGASNTGDGAAGTERKEGEVSGLDMNITQ
                     FLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQ
                     GTNPYLTFHCVSQGTILLDLAPDDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIR
                     IQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGM
                     FGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFST
                     MKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIVKPEAPSQTGTAA
                     EQKT"
ORIGIN

  Query Match           50.0%;  Score 1906.4;  DB 5;  Length 3804;
  Best Local Similarity 70.3%;  Pred. No. 0;
  Matches 2610;  Conservative    0;  Mismatches 1076;  Indels   28;  Gaps    3;

Qy          4 GCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCGTGG 63
              ||  | ||||||||||||  |||  | ||||     | ||  |  ||  |
Db         70 GCCTCGCCGCCCGCCGCCGCCTCTCCGCCGCGCAGCCCCAGCCTCGCGCCGGCCGAGCTC 129

Qy         64 GCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCC 123
              |  |   ||  |  ||  | | |  |    | |  ||||  |    | ||||   ||
Db        130 GGCCCCGCCGCGCAACGCCACAGCCTGGCCGGGCCCGAGGGCGAGGCTCCCCCCGACGCC 189

Qy        124 TGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCC 183
                 || ||    ||  | |  ||||     |  | | | |||||  |    |  ||||
Db        190 GAGCGGCCTCCGGCCCCGGAGTGCTCGGAGGGCGCCGCTCCGGGCCCGCCTCCGGGCTCG 249

Qy        184 GGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCG 243
              ||  || |   | |  ||| |  |         |   || |      |  | ||
Db        250 GGCAGCAGCTCCGGCAGCTCGGCCTCTTCCTCCTCCTCCTCGTCGTCCACGTCGTCGTCG 309

Qy        244 CCGGCG-GGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGA 302
                ||||  | |  ||||    | ||||  ||| ||| | || |  |     ||||  |||
Db        310 GTGGCGTCGAGCCCGGCGGCGGAGAGCCCCGAGGCGGCGGGCCCGAGCGGGGCCTTCCGC 369

Qy        303 GAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCT 362
              |||||| | ||||| |||| || |||||| || |||||| |||||     ||
Db        370 GAGCTGCTGGAGGCTTGCCGTAATGGAGACGTGACGCGCGTCAAGCGCCTGGTGGACGCT 429

Qy        363 GAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCA 422
              |  || || |||  |   |||| ||| ||| | || ||||| ||||| ||||||||||||
Db        430 GGCAACGTCAACGCCAAGGACATGGCCGGCCGCAAGTCCACGCCGCTCCACTTCGCCGCT 489

Qy        423 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 482
              || |||||  |||||||  || || ||  | |||||  ||||  ||  || || || || ||
Db        490 GGGTTTGGAAGGAAAGATGTTGTCGAGCACTTGCTACAGACAGGAGCCAACGTTCATGCT 549

Qy        483 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 542
              |||||||| || || || || ||||| || || |||||  |||||||||||||||||||
Db        550 CGTGATGACGGAGGGCTGATACCCCTTCACAACGCCTGCTCCTTTGGTCATGCTGAAGTG 609

Qy        543 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 602
              || | |||  | || | || ||||||||||| |||||  |||| |||||||| || |||
Db        610 GTTAGTCTGTTACTGTGCCAAGGTGCAGACCCTAATGCTAGAGACAATTGGAACTACACT 669

Qy        603 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 662
              ||| | || || ||||| || || || ||||| ||||  || |||||||| | |||||
Db        670 CCTTTGCACGAGGCTGCTATCAAGGGGAAGATAGATGTGTGTATTGTGCTGCTGCAGCAC 729

Qy        663 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 722
              |||||||| |||| ||| || || || ||||| |  | ||  |||||  | |||||  ||
Db        730 GGAGCTGATCCAAACATTCGGAACACCGATGGGAAATCTGCGCTGGATCTGGCAGACCCC 789

Qy        723 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 782
              || ||  |||| || || || |||||||| ||||| || ||||||  | |||  ||| |||
Db        790 TCAGCAGAAGCTGTACTCACAGGTGAATACAAGAAGGACGAACTCCTGGAAGCTGCTAGG 849

Qy        783 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 842
              ||||| |||||||||||| |||||||| || | || || |||||||| |||||||| |||
Db        850 AGTGGTAATGAAGAAAAACTGATGGCTTTATTGACTCCTTTAAATGTGAACTGCCATGCA 909

Qy        843 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 902
              || |||||  | |||||  |||||  |||||  | |||||||  || |||||||||  ||
Db        910 AGCGATGGACGCAAGTCCACTCCTTTACACCTAGCAGCAGGCTACAACAGAGTACGAATA 969

Qy        903 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 962
              || || ||  | || || ||||| || || || |||||||  ||| || ||  ||||| ||
Db        970 GTCCAACTTCTGCTTCAGCATGGTGCAGATGTTCACGCAAAGGACAAAGGTGGTCTTGTG 1029

Qy        963 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1022
              || | || || || |||||  ||||| || |||||||  || || ||  |  |||||
Db       1030 CCTCTTCATAACGCATGTTCATATGGACACTATGAAGTTACAGAGCTACTACTTAAGCAC 1089

Qy       1023 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1082
              |||||  ||  || |||||||||||| | ||||||| |||||  | ||||||||  |||||
Db       1090 GGTGCATGCGTTAATGCAATGGACCTCTGGCAATTTACTCCCTTGCATGAGGCTGCTTCC 1149

Qy       1083 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1142
              ||||||  || |||| |||||||  || |  ||| |||||||| || ||  | ||  ||
Db       1150 AAGAACCGTGTAGAAGTCTGTTCCCTCTTGCTGAGTCACGGTGCAGATCCCACGTTAGTC 1209

Qy       1143 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1202
              || || ||    || |||||  | ||  |||||||  |||||  || |||  || ||  |
Db       1210 AACTGCCATGGAAAGAGTGCGGTCGATATGGCTCCGACACCTGAGCTAAGGGAGAGGCTG 1269

Qy       1203 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1262
              | |||||||| |||| || || || ||||||||  || |||| || ||  |  |  |
Db       1270 ACCTATGAATTCAAAGGGCATTCTTTACTGCAAGCAGCCAGAGAGGCAGATTTAGCCAAA 1329

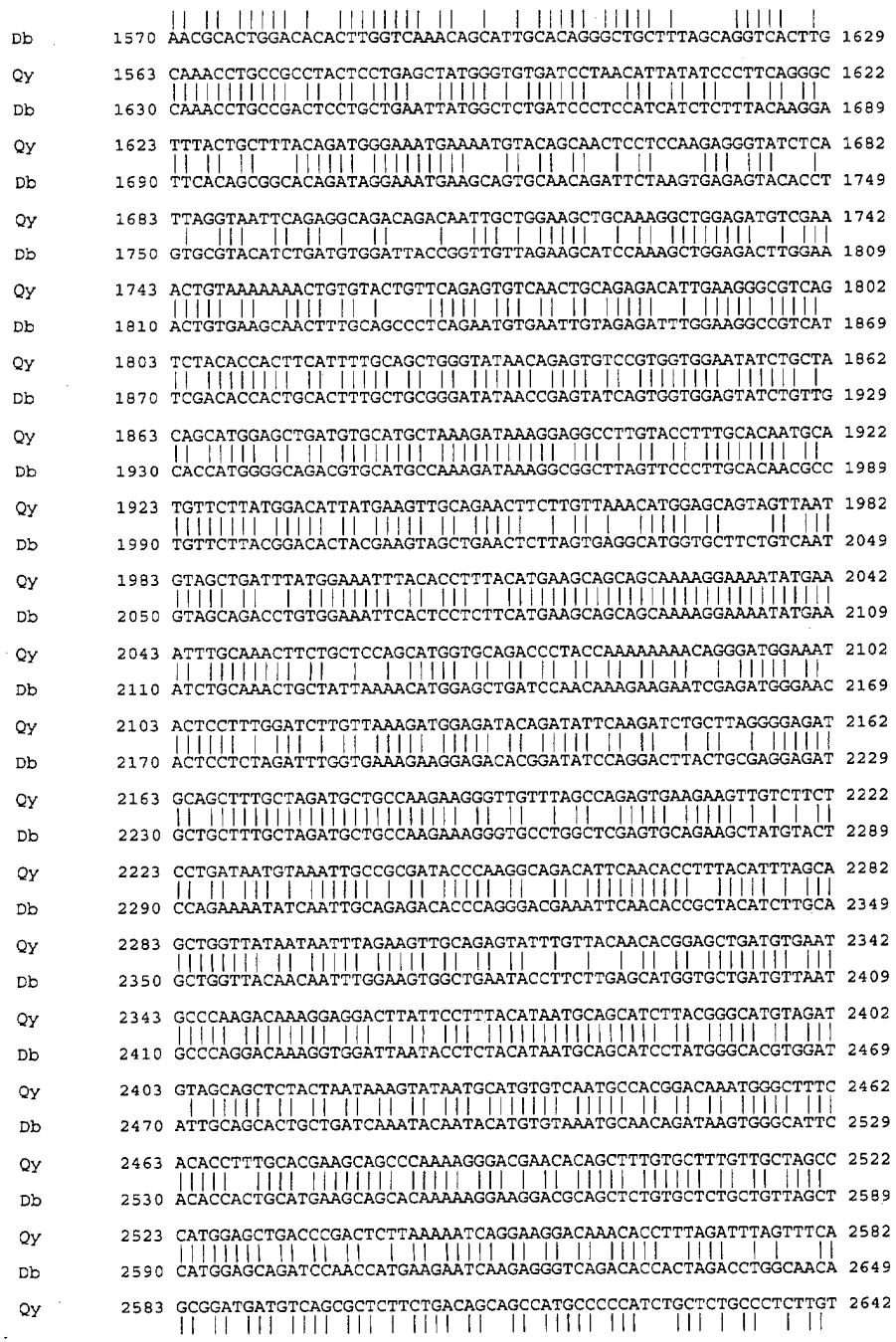
Qy       1263 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1322
              | ||||||   || |||| |  ||||| |||||  || ||  |||||| ||
Db       1330 GTGAAAAAAACACTTGCTTTGGAGATCATCAATTTTAAGCAGCCACAGTCACATGAGACT 1389

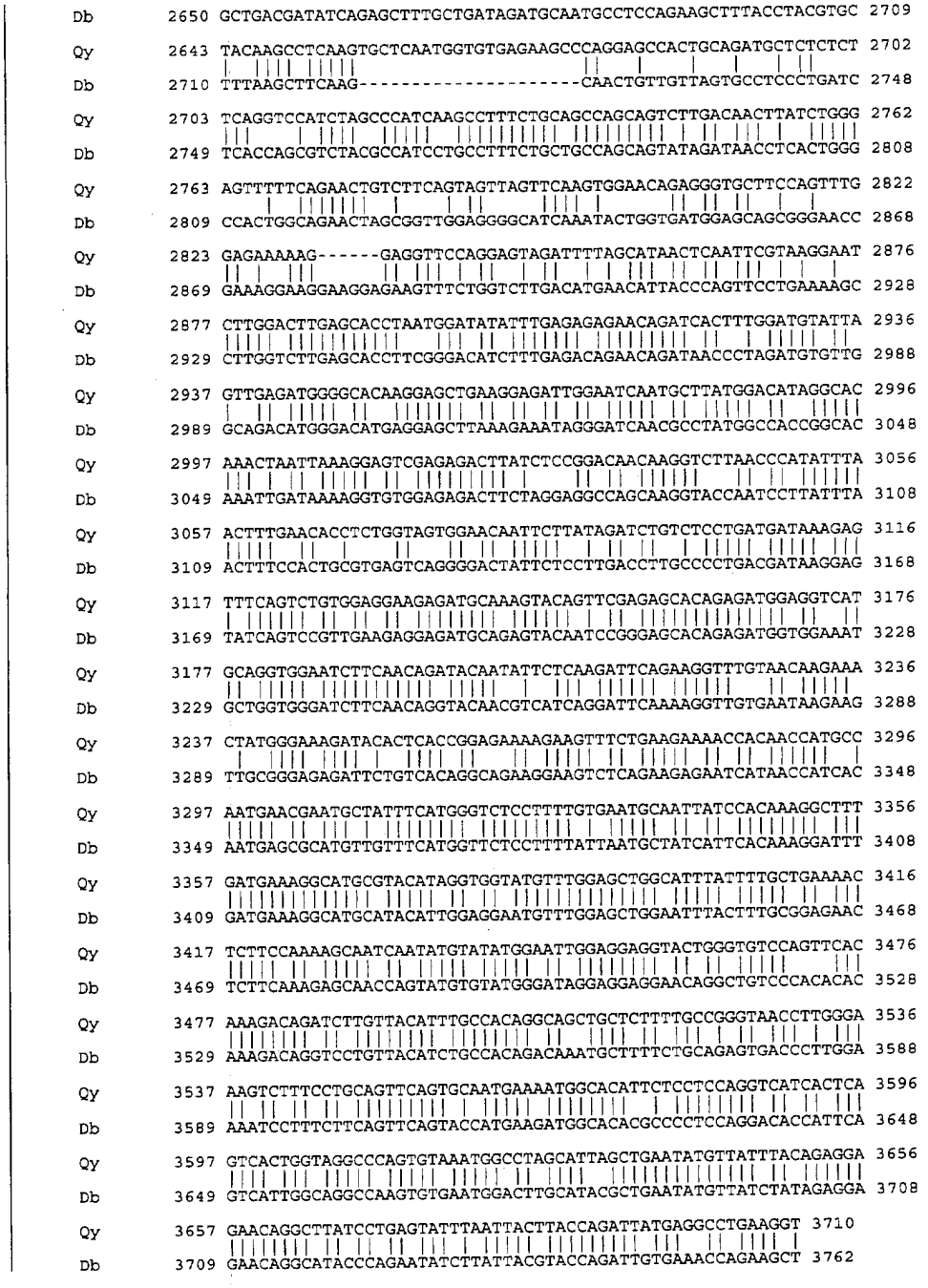
Qy       1323 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1382
              || |||| |||||||  || ||  |||||||  | |||||||  |  |||||  |||
Db       1390 GCTCTGCACTGTGCTGTGGCCGCTGTGCATCCCAAGCGGAAGCAAGTTACAGAACTATTG 1449

Qy       1383 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1442
              || || |||||||| |||| |||||| || |  |||| |||| |||| |||  |||| |||
Db       1450 CTCAGGAAAGGAGCCAACGTTAATGAGAAAAACAAAGATTTCATGACACCTTTGCATGTT 1509

Qy       1443 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1502
              ||  |||||||||||||||||||  | |||||  ||  |||| |||||||  |
Db       1510 GCCGCTGAGAAAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCATGGAGCAAAGATG 1569

Qy       1503 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1562
```


Db 1570 AACGCACTGGACACACTTGGTCAAACAGCATTGCACAGGGCTGCTTTAGCAGGTCACTTG 1629
Qy 1563 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1622
Db 1630 CAAACCTGCCGACTCCTGCTGAATTATGGCTCTGATCCCTCCATCATCTCTTTACAAGGA 1689
Qy 1623 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1682
Db 1690 TTCACAGCGGCACAGATAGGAAATGAAGCAGTGCAACAGATTCTAAGTGAGAGTACACCT 1749
Qy 1683 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1742
Db 1750 GTGCGTACATCTGATGTGGATTACCGGTTGTTAGAAGCATCCAAAGCTGGAGACTTGGAA 1809
Qy 1743 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1802
Db 1810 ACTGTGAAGCAACTTTGCAGCCCTCAGAATGTGAATTGTAGAGATTTGGAAGGCCGTCAT 1869
Qy 1803 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1862
Db 1870 TCGACACCACTGCACTTTGCTGCGGGATATAACCGAGTATCAGTGGTGGAGTATCTGTTG 1929
Qy 1863 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1922
Db 1930 CACCATGGGGCAGACGTGCATGCCAAAGATAAAGGCGGCTTAGTTCCCTTGCACAACGCC 1989
Qy 1923 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1982
Db 1990 TGTTCTTACGGACACTACGAAGTAGCTGAACTCTTAGTGAGGCATGGTGCTTCTGTCAAT 2049
Qy 1983 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2042
Db 2050 GTAGCAGACCTGTGGAAATTCACTCCTCTTCATGAAGCAGCAGCAAAAGGAAAATATGAA 2109
Qy 2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 2102
Db 2110 ATCTGCAAACTGCTATTAAAACATGGAGCTGATCCAACAAAGAAGAATCGAGATGGGAAC 2169
Qy 2103 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2162
Db 2170 ACTCCTCTAGATTTGGTGAAAGAAGGAGACACGGATATCCAGGACTTACTGCGAGGAGAT 2229
Qy 2163 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2222
Db 2230 GCTGCTTTGCTAGATGCTGCCAAGAAAGGGTGCCTGGCTCGAGTGCAGAAGCTATGTACT 2289
Qy 2223 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2282
Db 2290 CCAGAAAATATCAATTGCAGAGACACCCAGGGACGAAATTCAACACCGCTACATCTTGCA 2349
Qy 2283 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2342
Db 2350 GCTGGTTACAACAATTTGGAAGTGGCTGAATACCTTCTTGAGCATGGTGCTGATGTTAAT 2409
Qy 2343 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2402
Db 2410 GCCCAGGACAAAGGTGGATTAATACCTCTACATAATGCAGCATCCTATGGGCACGTGGAT 2469
Qy 2403 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2462
Db 2470 ATTGCAGCACTGCTGATCAAATACAATACATGTGTAAATGCAACAGATAAGTGGGCATTC 2529
Qy 2463 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2522
Db 2530 ACACCACTGCATGAAGCAGCACAAAAAGGAAGGACGCAGCTCTGTGCTCTGCTGTTAGCT 2589
Qy 2523 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2582
Db 2590 CATGGAGCAGATCCAACCATGAAGAATCAAGAGGGTCAGACACCACTAGACCTGGCAACA 2649
Qy 2583 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2642


Db 2650 GCTGACGATATCAGAGCTTTGCTGATAGATGCAATGCCTCCAGAAGCTTTACCTACGTGC 2709
Qy 2643 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2702
Db 2710 TTTAAGCTTCAAG----------CAACTGTTGTTAGTGCCTCCCTGATC 2748
Qy 2703 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2762
Db 2749 TCACCAGCGTCTACGCCATCCTGCCTTTCTGCTGCCAGCAGTATAGATAACCTCACTGGG 2808
Qy 2763 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2822
Db 2809 CCACTGGCAGAACTAGCGGTTGGAGGGGCATCAAATACTGGTGATGGAGCAGCGGGAACC 2868
Qy 2823 GAGAAAAAG------GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAAT 2876
Db 2869 GAAAGGAAGGAAGGAGAAGTTTCTGGTCTTGACATGAACATTACCCAGTTCCTGAAAAGC 2928
Qy 2877 CTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTA 2936
Db 2929 CTTGGTCTTGAGCACCTTCGGGACATCTTTGAGACAGAACAGATAACCCTAGATGTGTTG 2988
Qy 2937 GTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCAC 2996
Db 2989 GCAGACATGGGACATGAGGAGCTTAAAGAAATAGGGATCAACGCCTATGGCCACCGGCAC 3048
Qy 2997 AAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTA 3056
Db 3049 AAATTGATAAAAGGTGTGGAGAGACTTCTAGGAGGCCAGCAAGGTACCAATCCTTATTTA 3108
Qy 3057 ACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAG 3116
Db 3109 ACTTTCCACTGCGTGAGTCAGGGGACTATTCTCCTTGACCTTGCCCCTGACGATAAGGAG 3168
Qy 3117 TTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCAT 3176
Db 3169 TATCAGTCCGTTGAAGAGGAGATGCAGAGTACAATCCGGGAGCACAGAGATGGTGGAAAT 3228
Qy 3177 GCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAA 3236
Db 3229 GCTGGTGGGATCTTCAACAGGTACAACGTCATCAGGATTCAAAAGGTTGTGAATAAGAAG 3288
Qy 3237 CTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCC 3296
Db 3289 TTGCGGGAGAGATTCTGTCACAGGCAGAAGGAAGTCTCAGAAGAGAATCATAACCATCAC 3348
Qy 3297 AATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTT 3356
Db 3349 AATGAGCGCATGTTGTTTCATGGTTCTCCTTTTATTAATGCTATCATTCACAAAGGATTT 3408
Qy 3357 GATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAAC 3416
Db 3409 GATGAAAGGCATGCATACATTGGAGGAATGTTTGGAGCTGGAATTTACTTTGCGGAGAAC 3468
Qy 3417 TCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCAC 3476
Db 3469 TCTTCAAAGAGCAACCAGTATGTGTATGGGATAGGAGGAGGAACAGGCTGTCCCACACAC 3528
Qy 3477 AAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGA 3536
Db 3529 AAAGACAGGTCCTGTTACATCTGCCACAGACAAATGCTTTTCTGCAGAGTGACCCTTGGA 3588
Qy 3537 AAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCA 3596
Db 3589 AAATCCTTTCTTCAGTTCAGTACCATGAAGATGGCACACGCCCCTCCAGGACACCATTCA 3648
Qy 3597 GTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGA 3656
Db 3649 GTCATTGGCAGGCCAAGTGTGAATGGACTTGCATACGCTGAATATGTTATCTATAGAGGA 3708
Qy 3657 GAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
Db 3709 GAACAGGCATACCCAGAATATCTTATTACGTACCAGATTGTGAAACCAGAAGCT 3762

```
RESULT 35
AX062144
LOCUS       AX062144               3984 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 3 from Patent WO0100849.
ACCESSION   AX062144
VERSION     AX062144.1  GI:12540047
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 3 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3984
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .3984
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25053.1"
                     /db_xref="GI:12540048"
                     /db_xref="REMTREMBL:CAC25053"
                     /translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPPLSPGLAPGTT
                     PASPTASGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGTSCCSTTSTICTVAAA
                     PVVPAVSTSSAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPGSSLAESPEAAGVS
                     STAPLGPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSS
                     PLHFAAGFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPN
                     ARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEY
                     KKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHG
                     ADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEV
                     CSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKT
                     LALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
                     ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQG
                     FTAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEG
                     RHSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHG
                     ASVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQ
                     DLLKGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYL
                     LEHGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKG
                     RTQLCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVV
                     SASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNAGDGAAGTERKEGEVAGLDMN
                     ISQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLG
                     GQQGTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYN
                     VIRIQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYI
                     GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQ
                     FSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTA
                     TAAEQKT"
ORIGIN

  Query Match           49.4%;  Score 1885;  DB 6;  Length 3984;
  Best Local Similarity  72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy       252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db       502 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 561

Qy       312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db       562 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 621

Qy       372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db       622 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 681

Qy       432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db       682 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 741

Qy       492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db       742 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 801

Qy       552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db       802 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 861

Qy       612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db       862 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 921

Qy       672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db       922 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 981

Qy       732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db       982 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1041

Qy       792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db      1042 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1101

Qy       852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db      1102 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1161

Qy       912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db      1162 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1221

Qy       972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db      1222 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1281

Qy      1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db      1282 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1341

Qy      1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db      1342 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1401

Qy      1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db      1402 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1461

Qy      1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db      1462 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1521

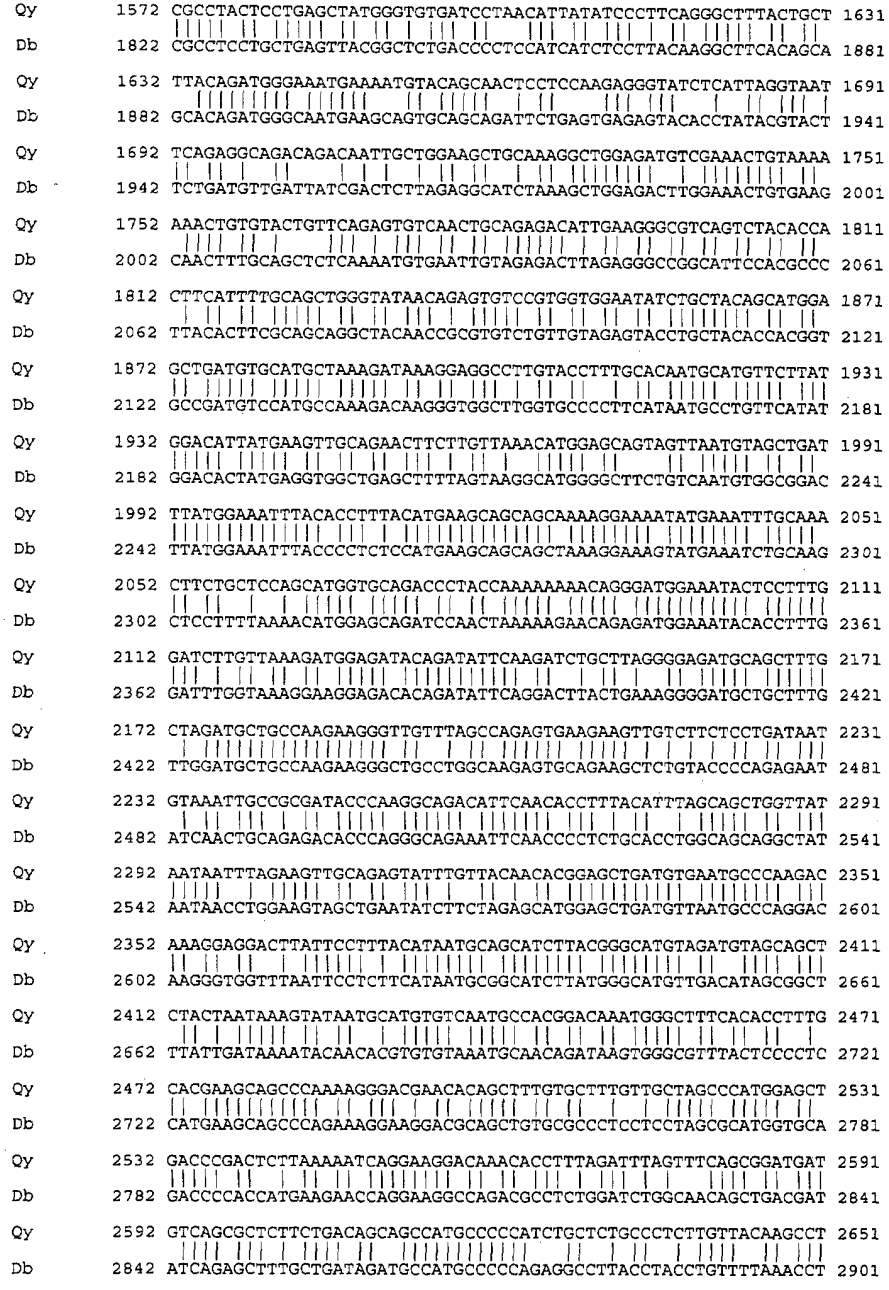
Qy      1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db      1522 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1581

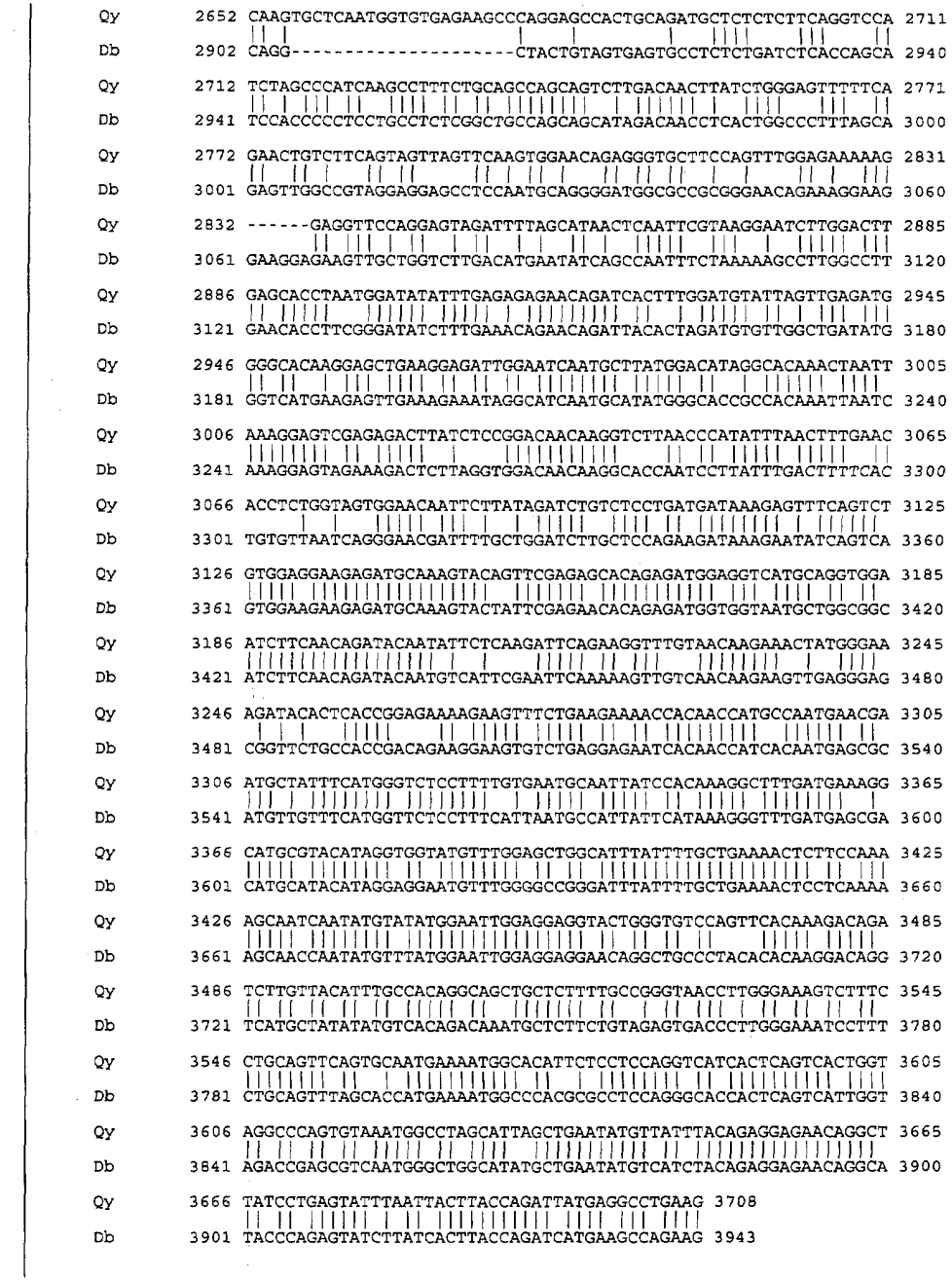
Qy      1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
Db      1582 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1641

Qy      1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db      1642 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1701

Qy      1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db      1702 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1761

Qy      1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db      1762 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1821
```





```
RESULT 36
BD235110
LOCUS       BD235110                4134 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Tankyrase, TRFI-binding protein.
ACCESSION   BD235110
VERSION     BD235110.1  GI:33044880
KEYWORDS    JP 2002517251-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4134)
  AUTHORS   Lange,T.D. and Smith,S.
  TITLE     Tankyrase, TRFI-binding protein
  JOURNAL   Patent: JP 2002517251-A 1 18-JUN-2002;
            THE ROCKEFELLER UNIVERSITY
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002517251-A/1
            PD   18-JUN-2002
            PF   09-JUN-1999 JP 2000553596
            PR   10-JUN-1998 US    09/095225,17-AUG-1998 US    09/135233 PR
            19-NOV-1998 US    09/196387
            PI   TITIA DE LANGE,SUSAN SMITH
            PC   C12N15/00,A01K67/027,A61K31/166,A61K38/45,A61K39/395,A61K39/
            PC   395,A61K45/00,
            PC   A61K48/00,A61P35/00,A61P43/00,C07K16/40,C12N9/12,C12Q1/48, PC
             G01N33/15,
            PC   G01N33/50,C12N15/00,A61K37/52
            CC   Tankyrase, TRFI-binding protein
            FH   Key             Location/Qualifiers
            FT   source          1. .4134
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .4134
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             49.4%;  Score 1885;  DB 6;  Length 4134;
  Best Local Similarity   72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | || |   || |||     ||||   || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              |||||  ||  |||||  |||||||||    || ||  ||||||||||||     |  |  || ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||        ||||  |||  |||  ||||  ||   ||  |||||||||||  ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              ||||  ||  ||  ||  |||  |  ||  ||  ||||   |||||  ||||||||  ||  ||||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              ||  ||  ||  ||  ||  ||||||||||  ||  ||||||||  ||||||||  ||  ||  |  |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              |  |||  |  ||  ||  ||  ||  ||  |||||   |||||  |||||  |||||  |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||  ||||||||  |||||  |||||  ||||||||||||  |  |||||  ||||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              ||||  |||  ||  ||  ||  |||||  |   ||||   ||||   |  ||||||||  ||  ||  |||

Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              ||  ||  |||||  ||||||||  ||||||||  |||||  |||||   |||  ||||||||  |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              ||||||||  |  ||||||  ||||  ||  ||   |||||||  ||  |||||  ||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              |||||||  |||||  ||||||  |  |||||  ||  ||  ||||||||     ||  ||  |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
              |  ||  ||  |||||  ||||||||  |||||  |||||  |||||||  ||  ||  |  ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              |||||  |||||  |||||  ||||||||||  ||  |||||   |  ||||||||  ||  |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              ||  |||||  |||||  |  |||||  ||  |||||  ||  ||  |||||  |||||  |  |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              ||  |||||  ||  |||  |  |  |||  ||||  ||  ||  ||  |  ||||  ||  ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
              |||||||||  |  ||  ||||||  ||  ||  ||  |  |  ||  |||||  |  ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              |||||||  ||  ||  ||  |||||  ||  |||||||  ||  |  ||  |  |||||||
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
              |||  |||||||||  |  ||||||||  ||  ||  ||  |  ||||||||||||  ||||
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
              |||||||  ||  ||||  |||||||  |  ||  |||  |  ||  ||||  ||  ||||||
Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
              |||||||  |  ||||||  |  ||||||  |||  ||||||||  |||||  ||  |||  |  ||
Db       1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
              |  |||  ||||||||||  |  |||||  ||  ||  ||||  ||  |||  |  |||||  |||
Db       1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

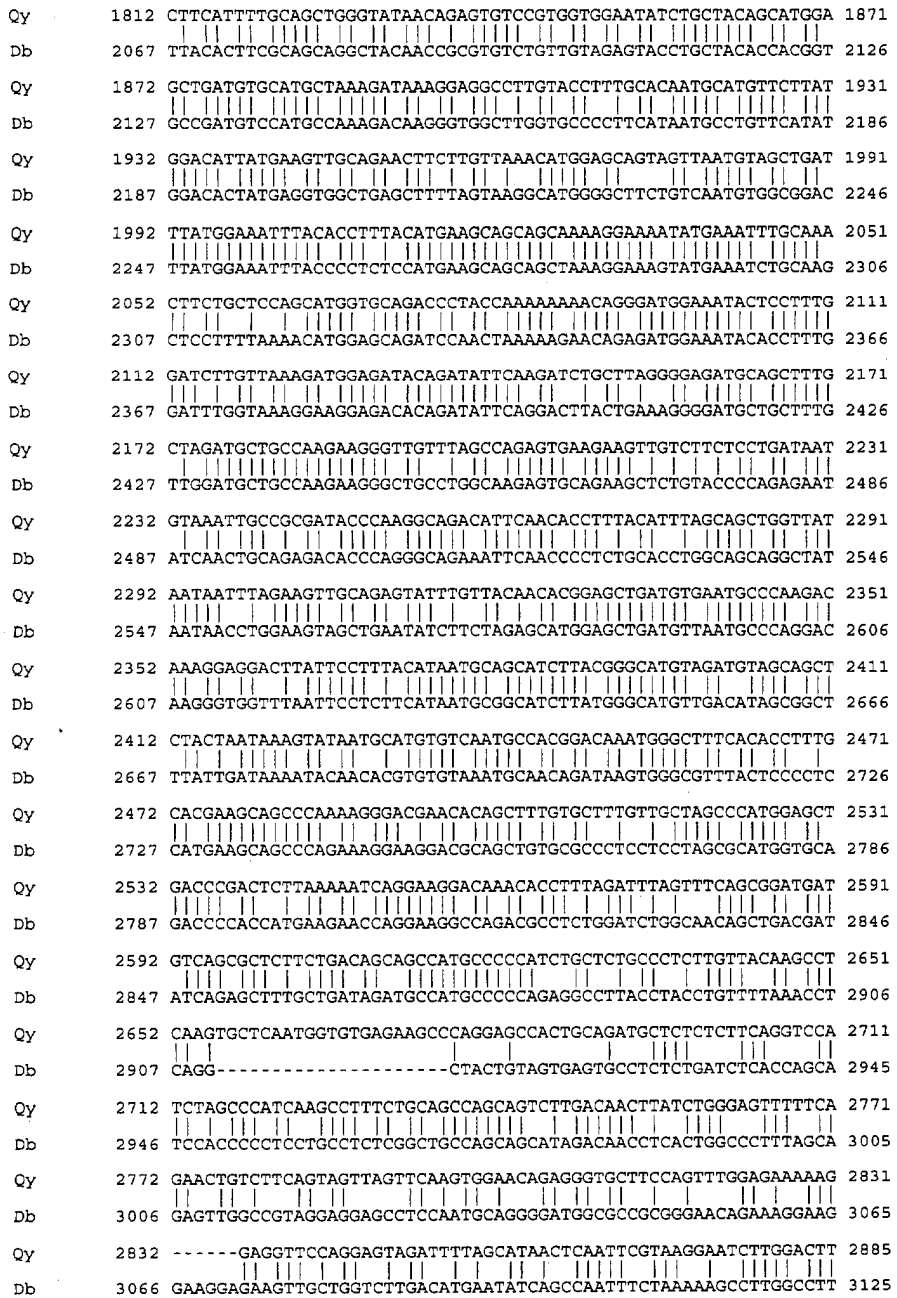
Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
              ||  |  ||||||||||||  ||  |  ||  |||||  ||  |||  ||  ||  ||||||||
Db       1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

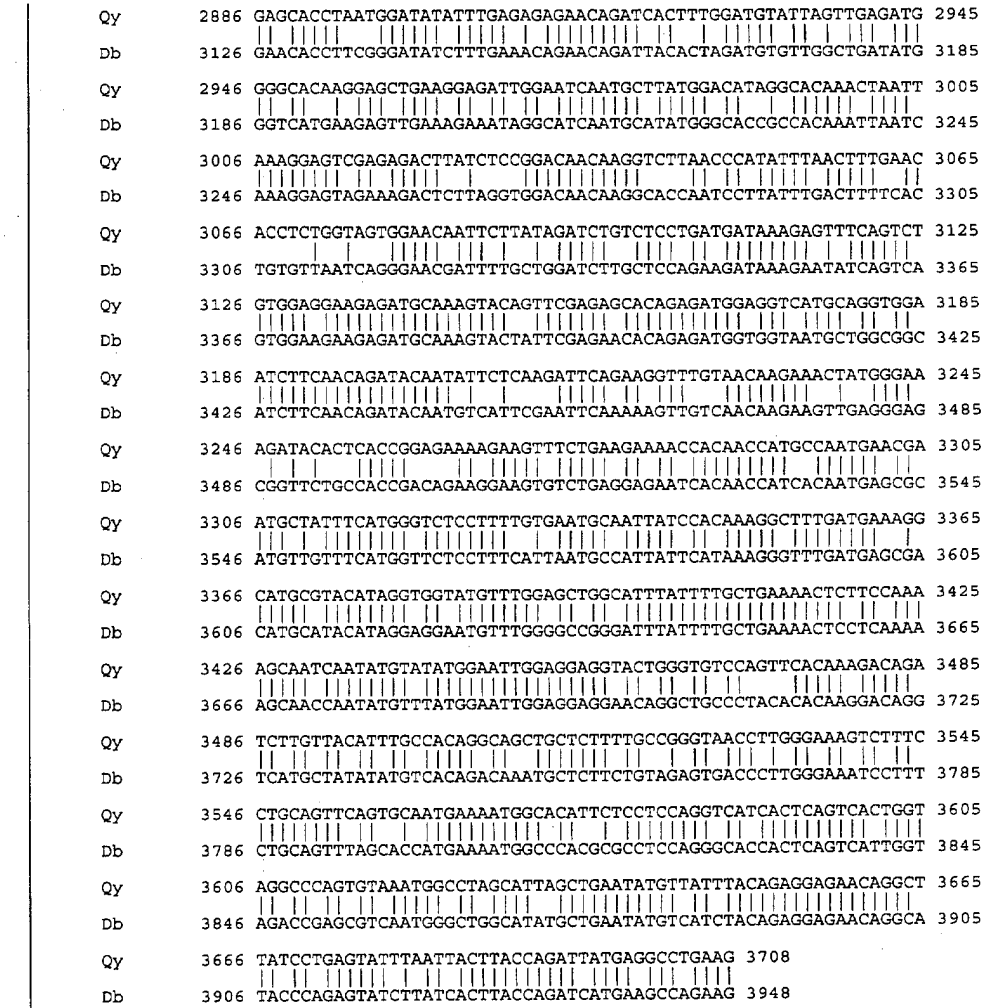
Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
              |||||  ||  |||||  ||  ||  |  |||  ||  |||  ||  |||  |  ||  |||||  ||  ||
Db       1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
                |||||||||  ||||||  ||  |||||  |  ||  |||  |||  |  ||  |||  |
Db       1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGAGAGTACACCTATACGTACT 1946

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAA 1751
              ||  ||  ||  |  |  |  ||  ||  |  ||  ||||||||  |  ||||||||  ||
Db       1947 TCTGATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAG 2006

Qy       1752 AAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCA 1811
              ||||  ||  |   |||  |  |||  ||  ||  ||||||  |  ||  ||  ||  ||  ||  ||
Db       2007 CAACTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCC 2066
```



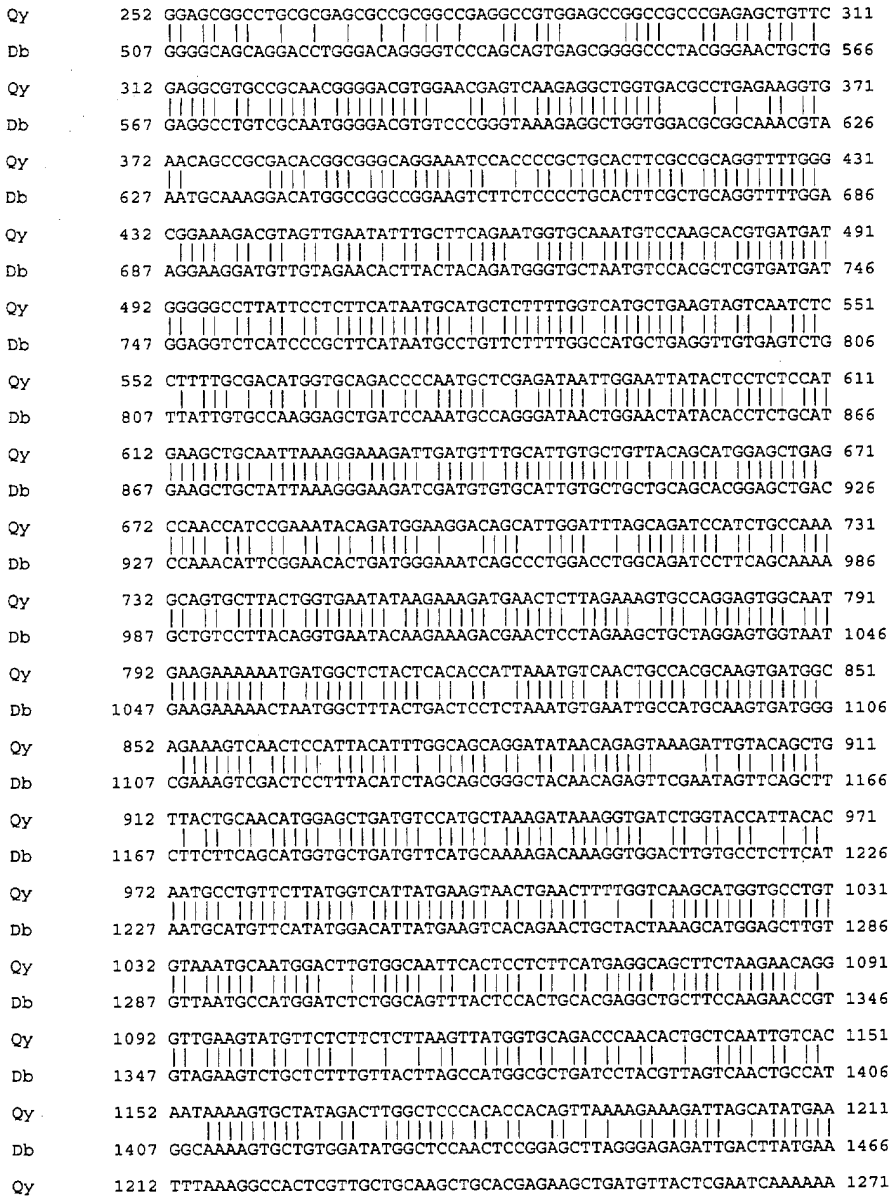


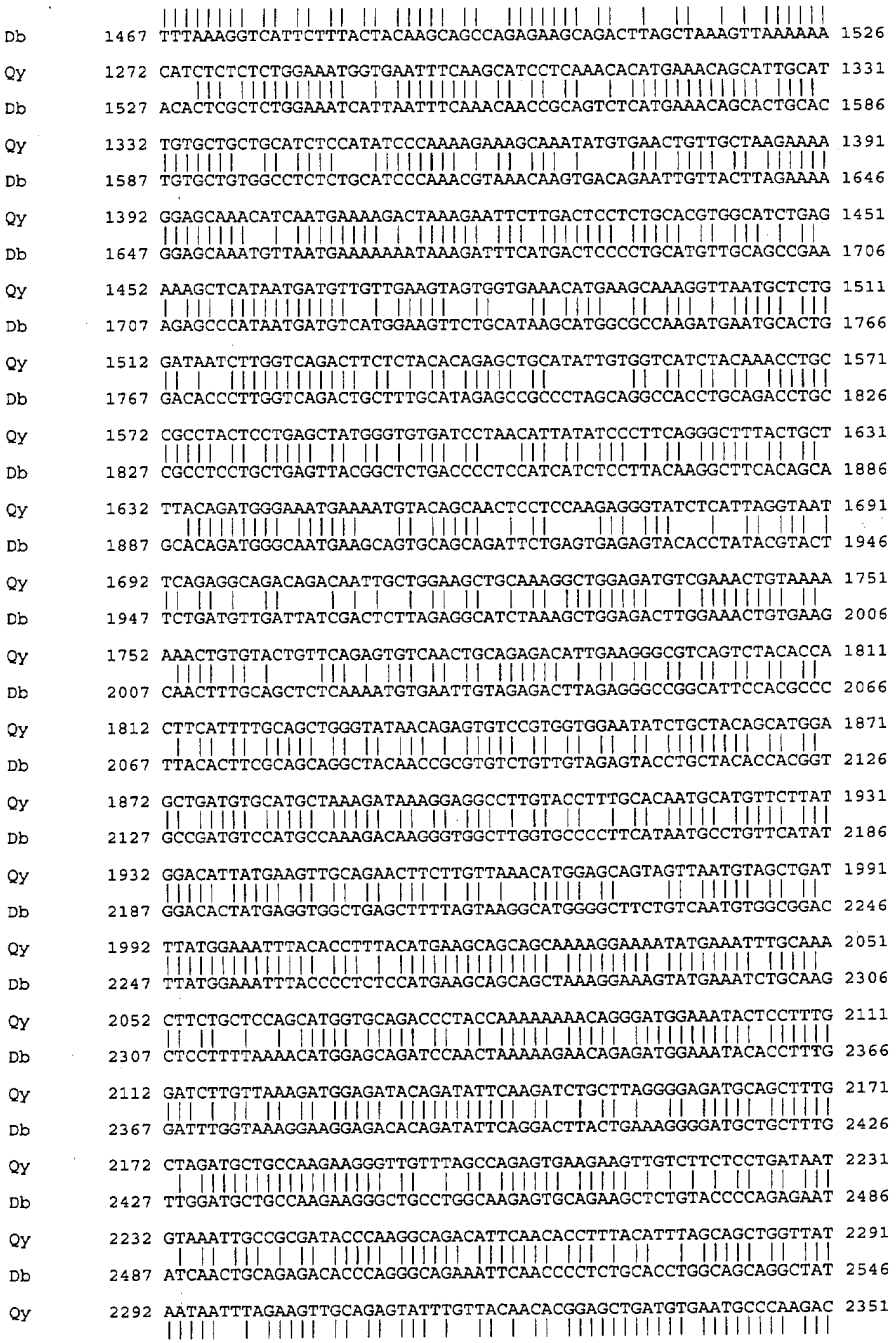
```
RESULT 37
AR274521
LOCUS       AR274521               4134 bp    mRNA    linear   PAT 10-APR-2003
DEFINITION  Sequence 1 from patent US 6506587.
ACCESSION   AR274521
VERSION     AR274521.1  GI:29707021
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4134)
  AUTHORS   De Lange,T. and Smith,S.
  TITLE     TRF 1 binding protein, methods of use thereof
  JOURNAL   Patent: US 6506587-A 1 14-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .4134
```

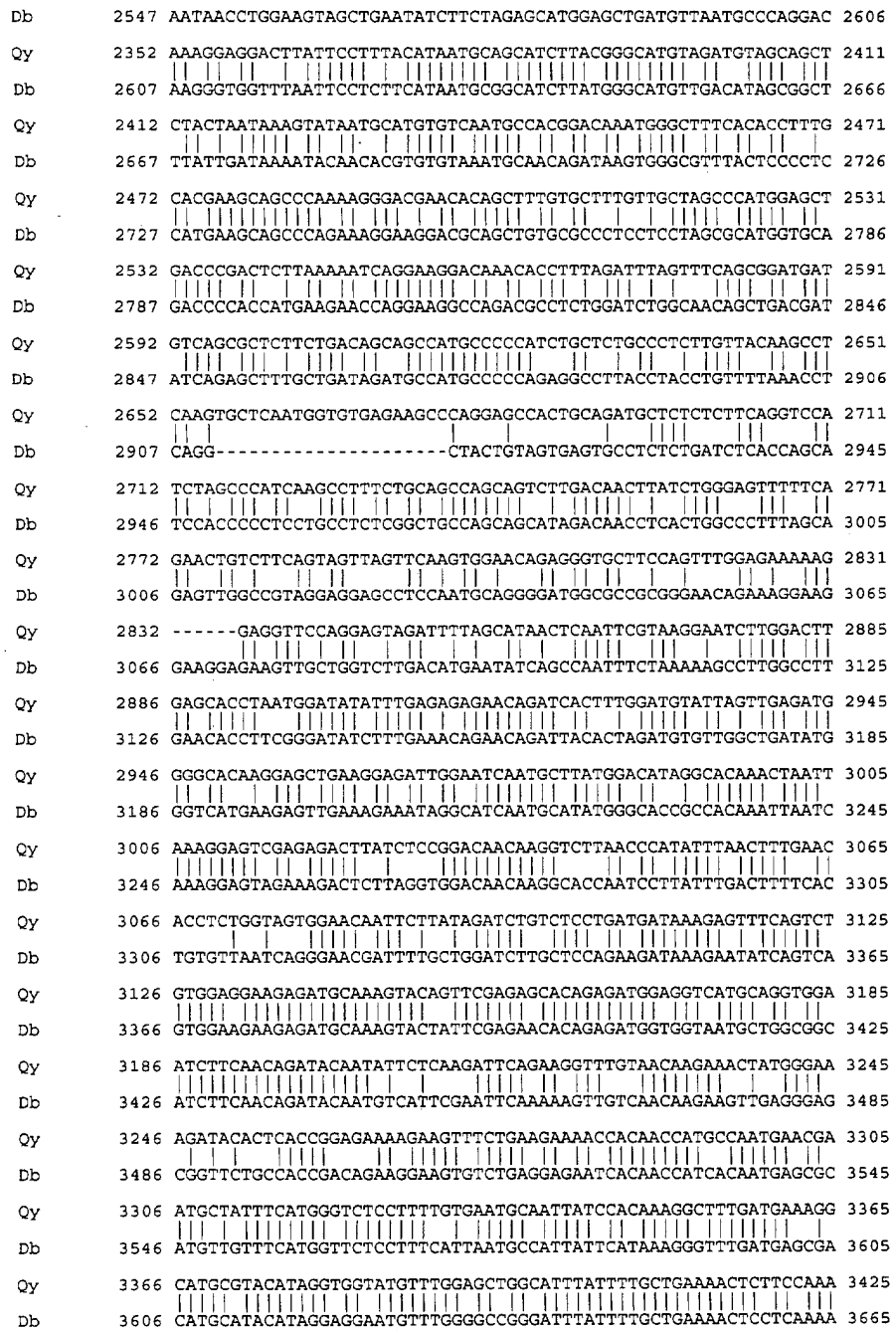
/organism="unknown"
/mol_type="mRNA"

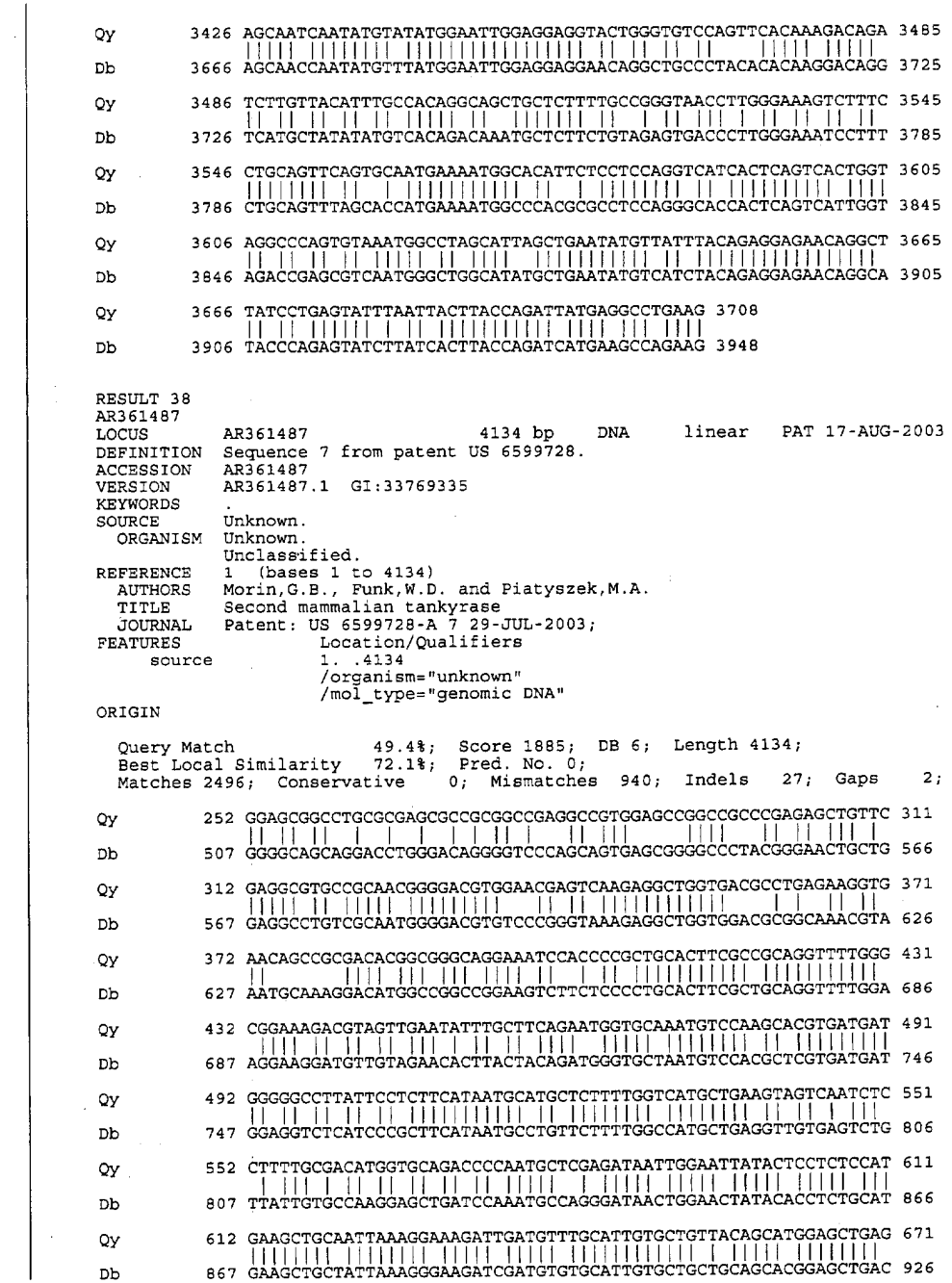
ORIGIN

Query Match 49.4%; Score 1885; DB 6; Length 4134;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps 2;



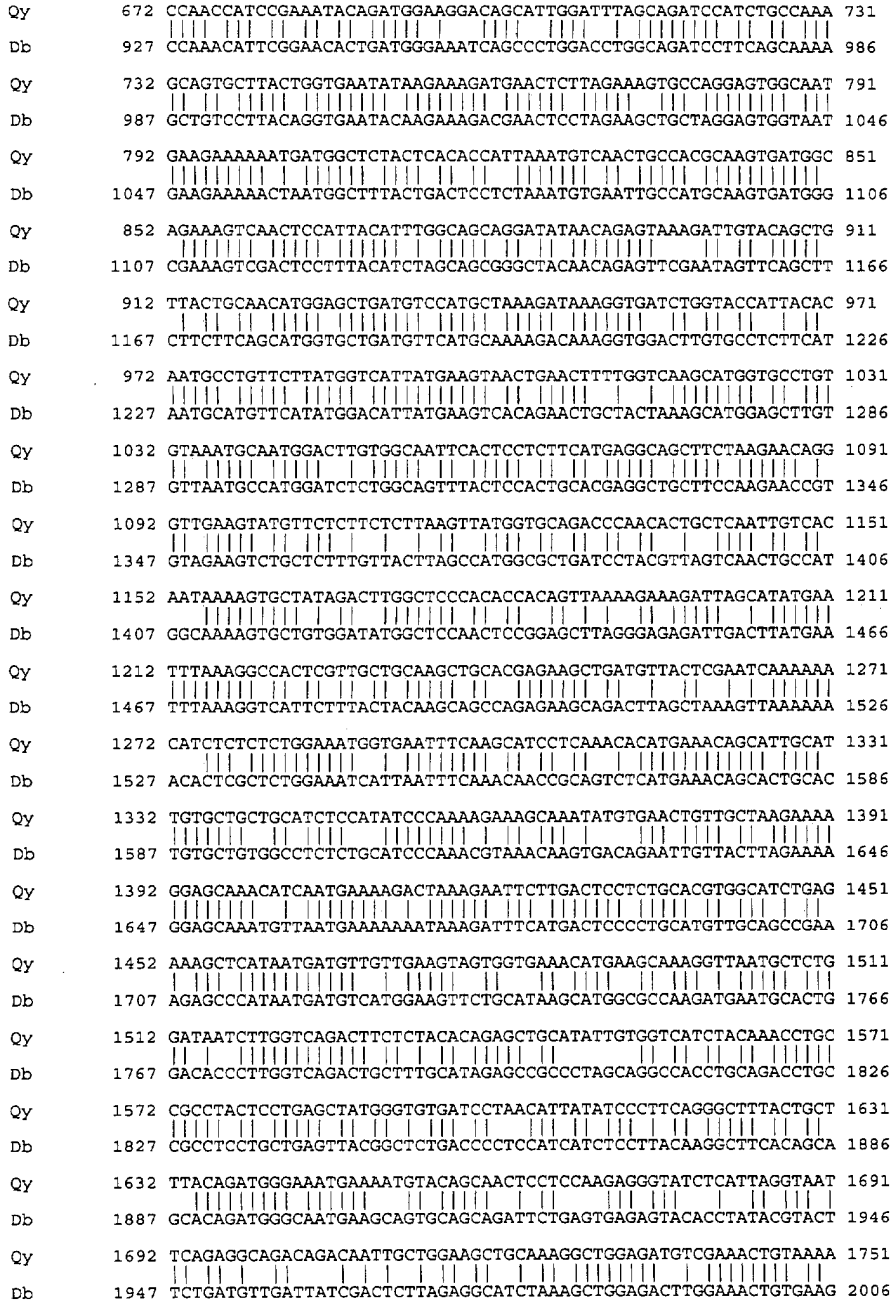


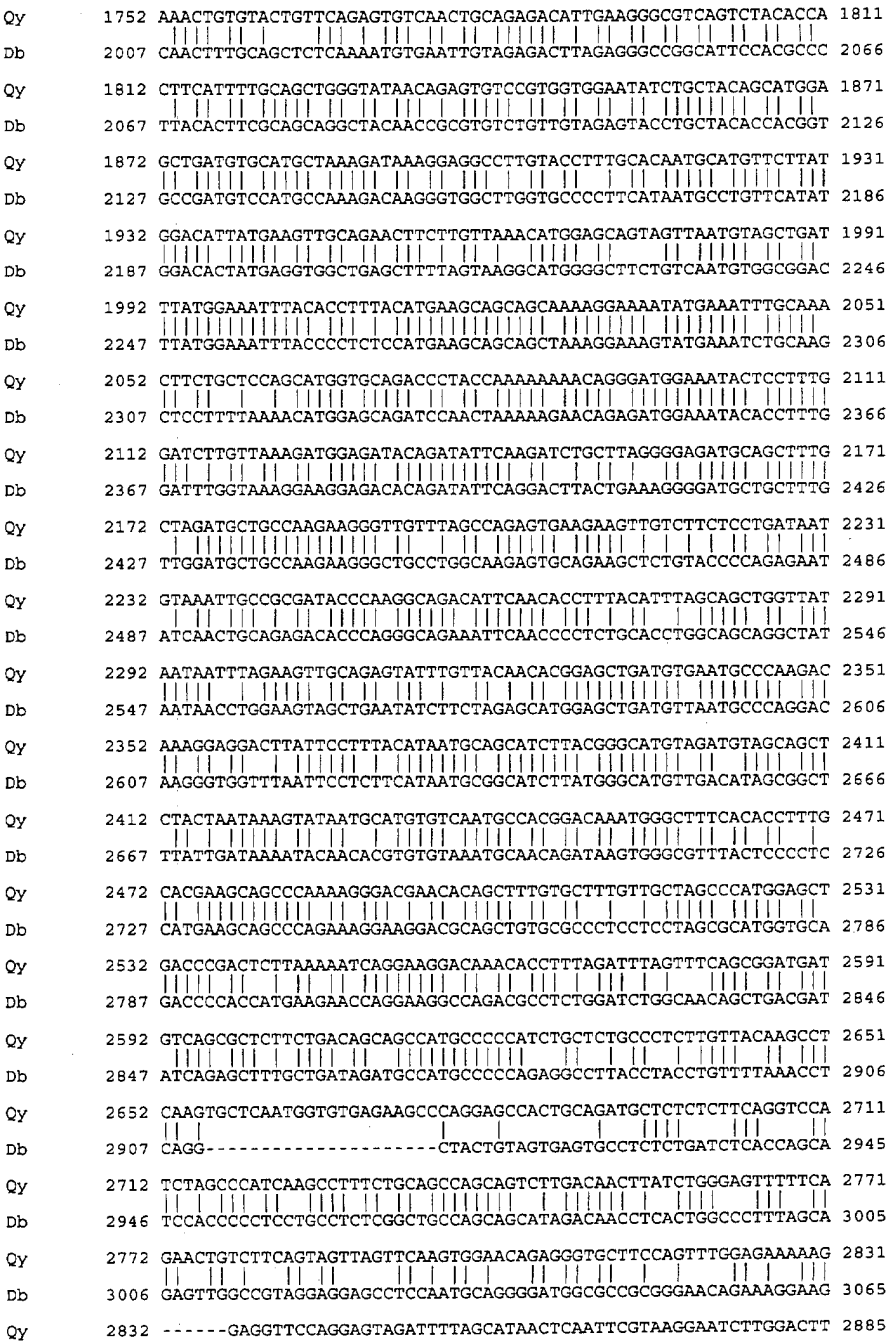




RESULT 38
AR361487
LOCUS AR361487 4134 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 7 from patent US 6599728.
ACCESSION AR361487
VERSION AR361487.1 GI:33769335
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4134)
AUTHORS Morin,G.B., Funk,W.D. and Piatyszek,M.A.
TITLE Second mammalian tankyrase
JOURNAL Patent: US 6599728-A 7 29-JUL-2003;
FEATURES Location/Qualifiers
source 1. .4134
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 49.4%; Score 1885; DB 6; Length 4134;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps 2;





```
Db       3066 GAAGGAGAAGTTGCTGGTCTTGACATGAATATCAGCCAATTTCTAAAAAGCCTTGGCCTT 3125

Qy       2886 GAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATG 2945

Db       3126 GAACACCTTCGGGATATCTTTGAAACAGAACAGATTACACTAGATGTGTTGGCTGATATG 3185

Qy       2946 GGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATT 3005

Db       3186 GGTCATGAAGAGTTGAAAGAAATAGGCATCAATGCATATGGGCACCGCCACAAATTAATC 3245

Qy       3006 AAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAAC 3065

Db       3246 AAAGGAGTAGAAAGACTCTTAGGTGGACAACAAGGCACCAATCCTTATTTGACTTTTCAC 3305

Qy       3066 ACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCT 3125

Db       3306 TGTGTTAATCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCA 3365

Qy       3126 GTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGA 3185

Db       3366 GTGGAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGC 3425

Qy       3186 ATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAA 3245

Db       3426 ATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAG 3485

Qy       3246 AGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGA 3305

Db       3486 CGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGC 3545

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365

Db       3546 ATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGA 3605

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425

Db       3606 CATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAA 3665

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485

Db       3666 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3725

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545

Db       3726 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3785

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605

Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665

Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708

Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948
```

```
RESULT 39
AX467061
LOCUS       AX467061                4134 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 1 from Patent WO0229017.
ACCESSION   AX467061
VERSION     AX467061.1  GI:21900390
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   de Lange,T. and Smith,S.
  TITLE     A novel trf1 binding protein, methods of use thereof
  JOURNAL   Patent: WO 0229017-A 1 11-APR-2002;
            THE ROCKEFELLER UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1. .4134
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="cDNA"
ORIGIN
```

```
  Query Match             49.4%;  Score 1885;  DB 6;  Length 4134;
  Best Local Similarity   72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311

Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371

Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431

Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491

Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551

Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611

Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671

Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731

Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791

Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851

Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911

Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971

Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

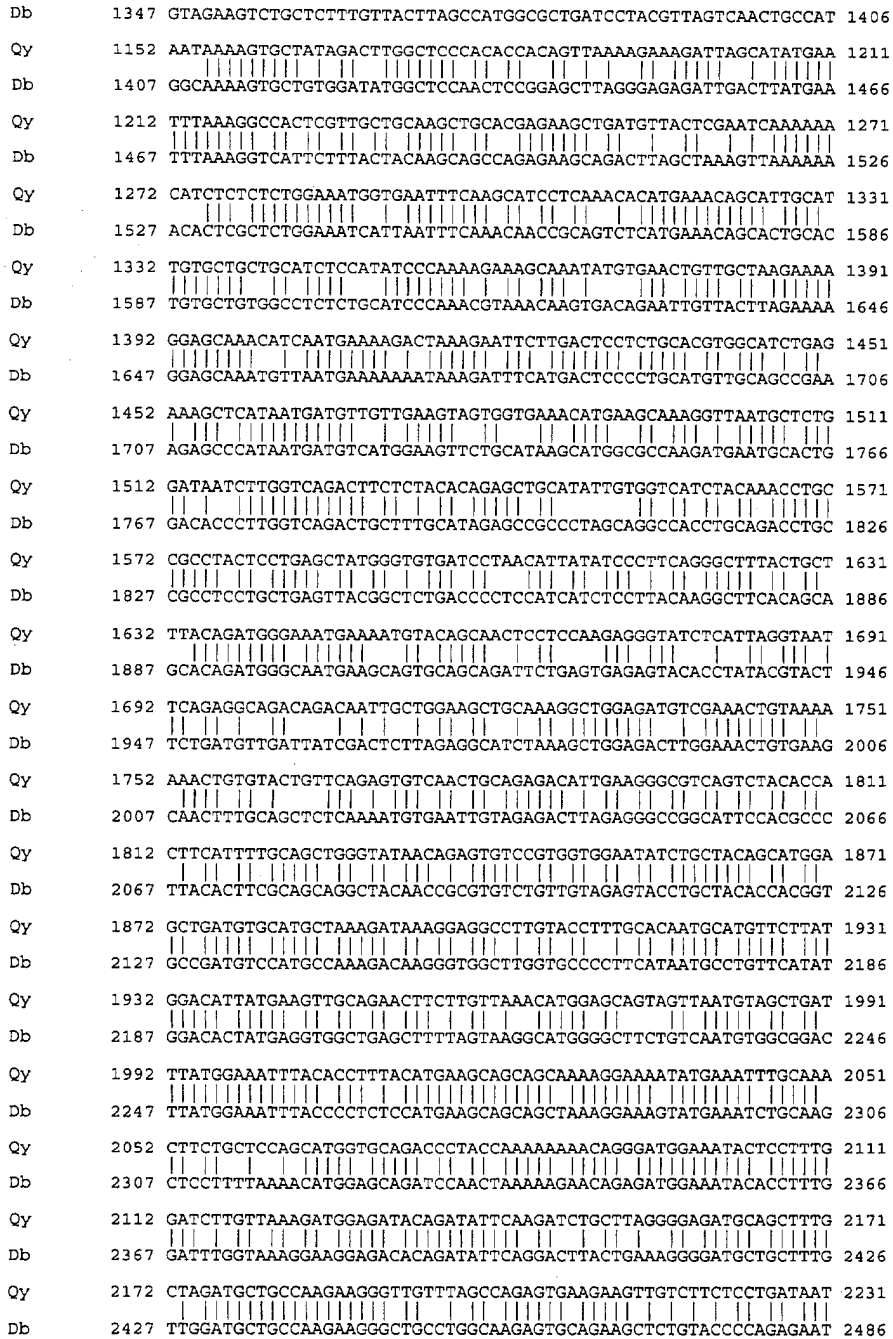
Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031

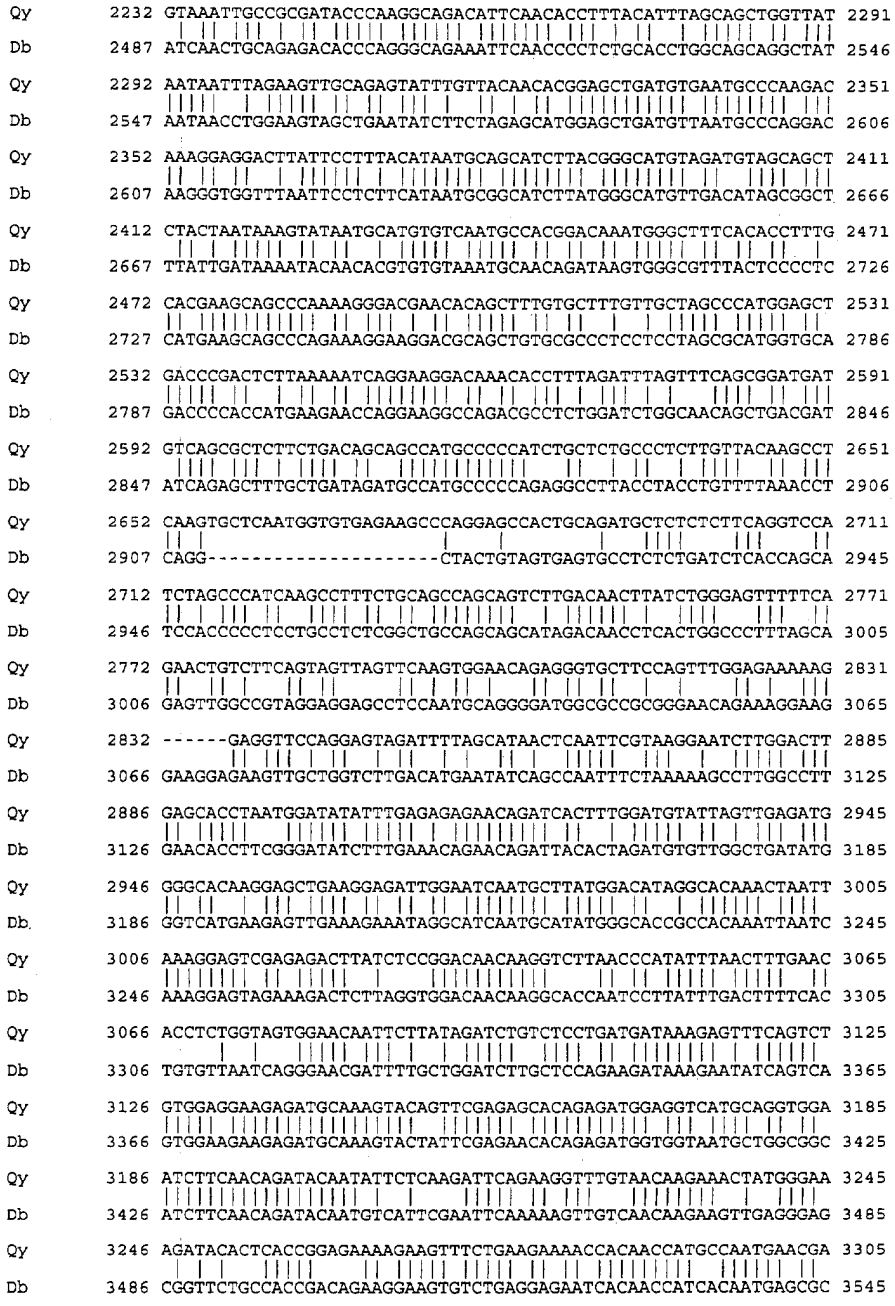
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091

Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
```





```
Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365
Db       3546 ATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGA 3605

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425
Db       3606 CATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAA 3665

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485
Db       3666 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3725

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545
Db       3726 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3785

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948
```

```
RESULT 40
AF082556
LOCUS       AF082556                4134 bp    mRNA    linear   PRI 30-NOV-1998
DEFINITION  Homo sapiens TRF1-interacting ankyrin-related ADP-ribose polymerase
            mRNA, complete cds.
ACCESSION   AF082556
VERSION     AF082556.1  GI:3929218
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4134)
  AUTHORS   Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
  TITLE     Tankyrase, a poly(ADP-ribose) polymerase at human telomeres
  JOURNAL   Science 282 (5393), 1484-1487 (1998)
  MEDLINE   99040105
   PUBMED   9822378
REFERENCE   2  (bases 1 to 4134)
  AUTHORS   Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1998) Laboratory for Cell Biology and Genetics,
            The Rockefeller University, 1230 York Avenue, New York, NY 10021,
            USA
FEATURES             Location/Qualifiers
     source          1. .4134
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
     CDS             6. .3989
                     /note="tankyrase"
                     /codon_start=1
                     /product="TRF1-interacting ankyrin-related ADP-ribose
                     polymerase"
                     /protein_id="AAC79841.1"
                     /db_xref="GI:3929219"
                     /translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPPLSPGLAPGTT
                     PASPTASGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGTSCCSTTSTICTVAAA
                     PVVPAVSTSSAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPGSSLAESPEAAGVS
                     STAPLGPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSS
                     PLHFAAGFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPN
                     ARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEY
                     KKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHG
                     ADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEV
                     CSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKT
                     LALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
                     ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQG
                     FTAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEG
                     RHSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHG
                     ASVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQ
                     DLLKGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYL
                     LEHGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKG
                     RTQLCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVV
                     SASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNAGDGAAGTERKEGEVAGLDMN
                     ISQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLG
                     GQQGTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYN
                     VIRIQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGPDERHAYI
                     GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQ
                     FSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTA
                     TAAEQKT"
ORIGIN

  Query Match           49.4%;  Score 1885;  DB 9;  Length 4134;
  Best Local Similarity  72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

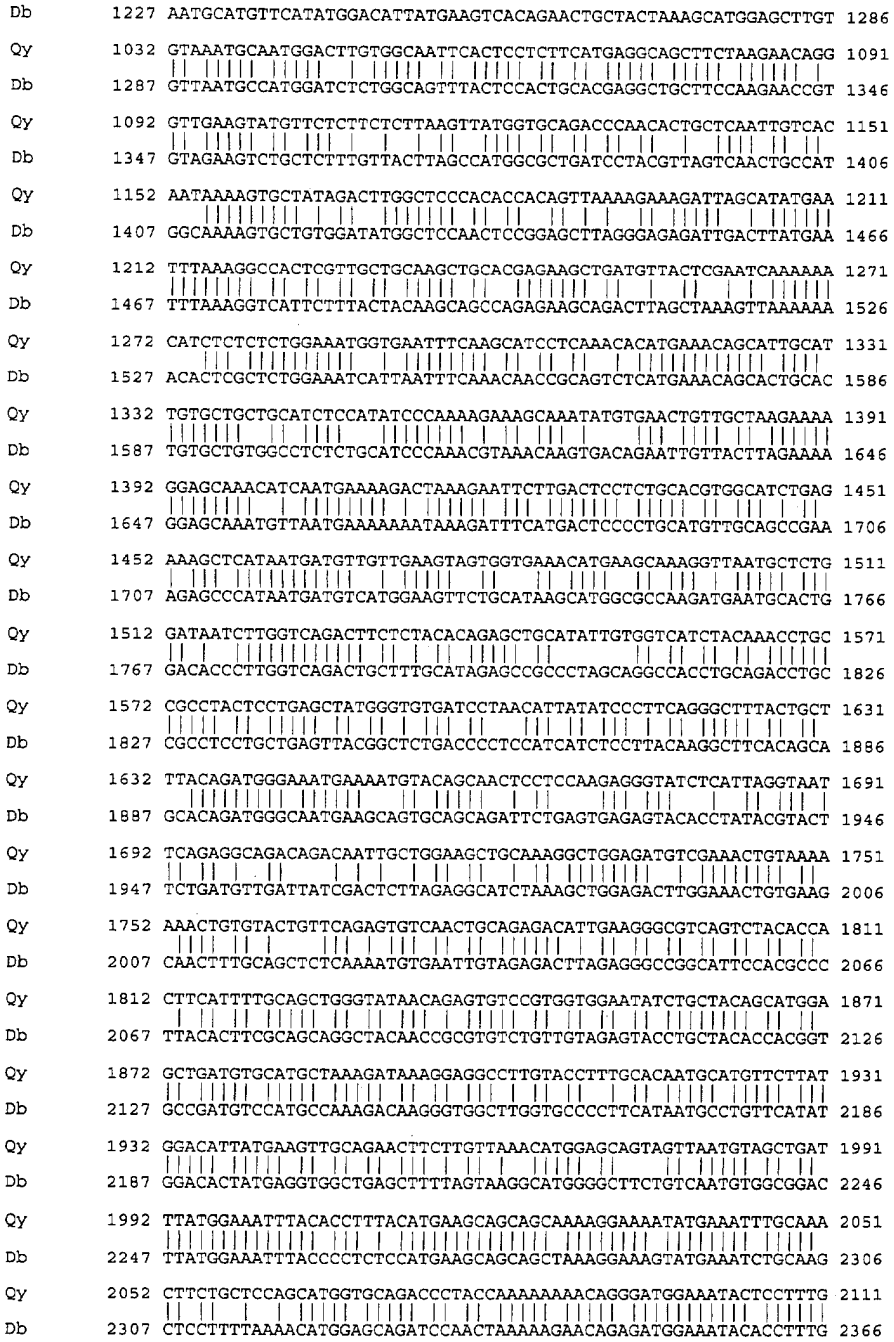
Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

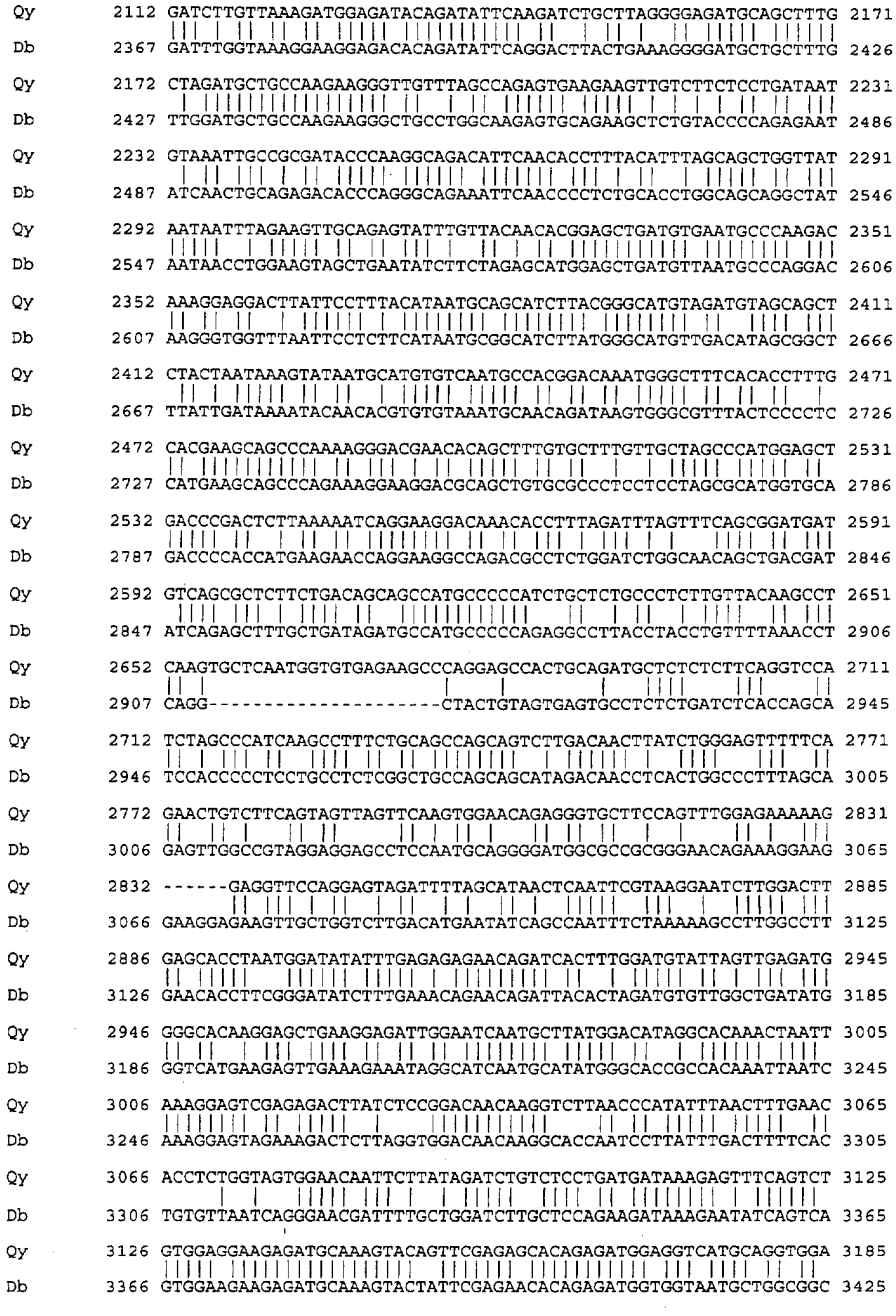
Qy        792 GAAGAAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
```





```
Qy       3186 ATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAA 3245
Db       3426 ATCTTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAG 3485

Qy       3246 AGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGA 3305
Db       3486 CGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGC 3545

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365
Db       3546 ATGTTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGA 3605

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425
Db       3606 CATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAA 3665

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485
Db       3666 AGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGG 3725

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545
Db       3726 TCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTT 3785

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
Db       3786 CTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGT 3845

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
Db       3846 AGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCA 3905

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
Db       3906 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3948
```

```
RESULT 41
AF082557
LOCUS       AF082557               8901 bp    mRNA    linear   PRI 30-NOV-1998
DEFINITION  Homo sapiens TRF1-interacting ankyrin-related ADP-ribose polymerase
            mRNA, partial cds.
ACCESSION   AF082557
VERSION     AF082557.1  GI:3929220
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8901)
  AUTHORS   Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
  TITLE     Tankyrase, a poly(ADP-ribose) polymerase at human telomeres
  JOURNAL   Science 282 (5393), 1484-1487 (1998)
  MEDLINE   99040105
   PUBMED   9822378
REFERENCE   2  (bases 1 to 8901)
  AUTHORS   Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1998) Laboratory for Cell Biology and Genetics,
            The Rockefeller University, 1230 York Avenue, New York, NY 10021,
            USA
FEATURES             Location/Qualifiers
     source          1. .8901
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="HeLa"
     CDS             <1. .3285
                     /note="tankyrase"
                     /codon_start=1
                     /product="TRF1-interacting ankyrin-related ADP-ribose
                     polymerase"
                     /protein_id="AAC79842.1"
                     /db_xref="GI:3929221"
                     /translation="HLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNA
                     RDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYK
                     KDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGA
                     DVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVC
                     SLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTL
                     ALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAE
                     RAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGF
                     TAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGR
                     HSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGA
                     SVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQD
                     LLKGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLL
                     EHGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGR
                     TQLCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVVS
                     ASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNAGDGAAGTERKEGEVAGLDMNI
                     SQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGG
                     QQGTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNV
                     IRIQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIG
                     GMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQF
                     STMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTAT
                     AAEQKT"
ORIGIN
```

```
  Query Match           47.1%;  Score 1796.4;  DB 9;  Length 8901;
  Best Local Similarity  72.5%;  Pred. No. 0;
  Matches 2366;  Conservative    0;  Mismatches  871;  Indels   27;  Gaps    2;

Qy        451 ATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTC 510
Db          2 ACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTC 61

Qy        511 ATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAG 570
Db         62 ATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTG 121

Qy        571 ACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAA 630
Db        122 ATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGA 181

Qy        631 AGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAG 690
Db        182 AGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTG 241

Qy        691 ATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAAT 750
Db        242 ATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAAT 301

Qy        751 ATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTC 810
Db        302 ACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTT 361

Qy        811 TACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTAC 870
Db        362 TACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTAC 421

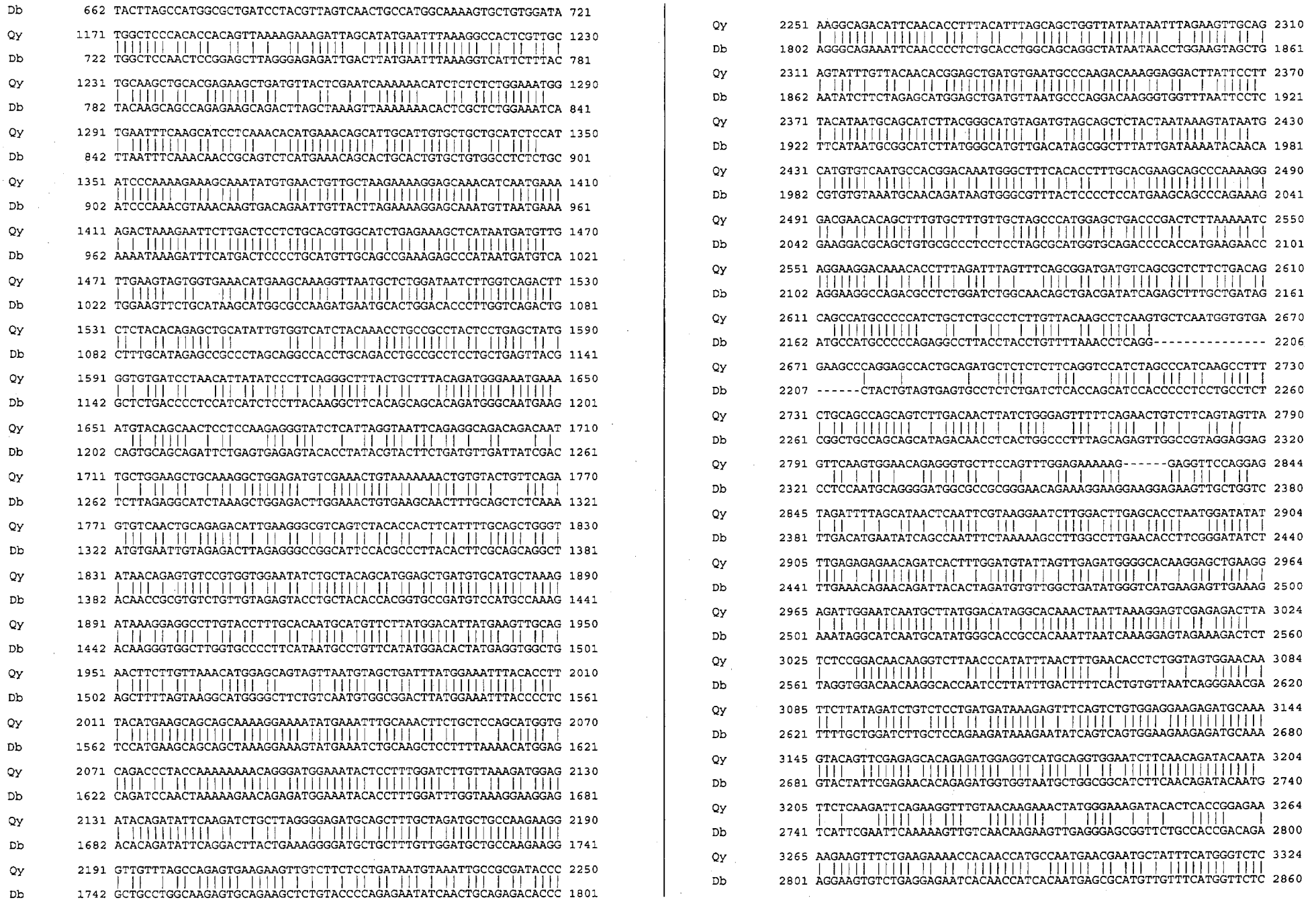
Qy        871 ATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTG 930
Db        422 ATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTG 481

Qy        931 ATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTC 990
Db        482 ATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGAC 541

Qy        991 ATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGT 1050
Db        542 ATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCT 601

Qy       1051 GGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTC 1110
Db        602 GGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGT 661

Qy       1111 TCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACT 1170
```



```
Qy       3325 CTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTA 3384
Db       2861 CTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAA 2920

Qy       3385 TGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATG 3444
Db       2921 TGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATG 2980

Qy       3445 GAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACA 3504
Db       2981 GAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACA 3040

Qy       3505 GGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGA 3564
Db       3041 GACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGA 3100

Qy       3565 AAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCC 3624
Db       3101 AAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGC 3160

Qy       3625 TAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTA 3684
Db       3161 TGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCA 3220

Qy       3685 CTTACCAGATTATGAGGCCTGAAG 3708
Db       3221 CTTACCAGATCATGAAGCCAGAAG 3244
```

```
RESULT 42
AX062239
LOCUS       AX062239               1568 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 98 from Patent WO0100849.
ACCESSION   AX062239
VERSION     AX062239.1  GI:12540140
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 98 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1568
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .>1567
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25059.1"
                     /db_xref="GI:12540141"
                     /db_xref="REMTREMBL:CAC25059"
                     /translation="ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVK
                     RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNAC
                     SFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG
                     RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPL
                     HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM
                     DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFK
                     GHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK
                     GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ
                     TCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV
                     ETVKKLCTVQSVNC"
ORIGIN

  Query Match             41.1%;  Score 1568;  DB 6;  Length 1568;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-294;
  Matches 1568;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274
```

```
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600

Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840

Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900

Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
Db        901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960

Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
Db        961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020

Qy       1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
Db       1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080

Qy       1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
```

```
Db       1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140

Qy       1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200

Qy       1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260

Qy       1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320

Qy       1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380

Qy       1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440

Qy       1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500

Qy       1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560

Qy       1775 CAACTGCA 1782
             ||||||||
Db       1561 CAACTGCA 1568


RESULT 43
BD235115
LOCUS       BD235115                4491 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Tankyrase, TRFI-binding protein.
ACCESSION   BD235115
VERSION     BD235115.1  GI:33044885
KEYWORDS    JP 2002517251-A/6.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4491)
  AUTHORS   Lange,T.D. and Smith,S.
  TITLE     Tankyrase, TRFI-binding protein
  JOURNAL   Patent: JP 2002517251-A 6 18-JUN-2002;
            THE ROCKEFELLER UNIVERSITY
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002517251-A/6
            PD   18-JUN-2002
            PF   09-JUN-1999 JP 2000553596
            PR   10-JUN-1998 US    09/095225,17-AUG-1998 US    09/135233 PR
            19-NOV-1998 US    09/196387
            PI   TITIA DE LANGE,SUSAN SMITH
            PC   C12N15/00,A01K67/027,A61K31/166,A61K38/45,A61K39/395,A61K39/
            PC   395,A61K45/00,
            PC   A61K48/00,A61P35/00,A61P43/00,C07K16/40,C12N9/12,C12Q1/48, PC
             G01N33/15,
            PC   G01N33/50,C12N15/00,A61K37/52
            CC   Description of Artificial Sequence: Primer
            CC   Human Truncated
            FH   Key             Location/Qualifiers
            FT   source          1. .4491
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .4491
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"

ORIGIN

  Query Match           39.9%;  Score 1521.2;  DB 6;  Length 4491;
  Best Local Similarity   65.4%;  Pred. No. 7.5e-285;
  Matches 2498;  Conservative    0;  Mismatches  938;  Indels  384;  Gaps    3;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
              || || ||  |  |  |  | | ||  || |||     ||||   || || ||| |
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
              |||||  || |||||  |||||||||   || ||  ||||||||||||    |  |  ||  ||
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
              ||          |||| |||  |||  |||| ||   ||  |||||||||||  ||||||||||
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
              ||||  ||  ||  ||  |||  |  ||  ||  ||||   |||||  ||||||||  ||  ||||||||||
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
              ||  ||  ||  ||  ||  ||||||||||  ||  ||||||||  ||||||||  ||  ||  |  |||
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
              |  |||  |  ||  ||  ||  ||  ||  |||||   |  |||||  |||||  |||||  |||||  |||
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
              |||||||  |||||||  |||||  |||||  ||||||||||  |  |||||  |||||||
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
              ||||  |||  ||  ||  ||  |||||  |    ||||  ||||   |  ||||||||  ||  ||  |||
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
              ||  ||  |||||  ||||||||  ||||||||  ||||||  |||||   |||  ||||||||  |||
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
              |||||||||  |  ||||||  ||||  ||  ||  |||||||  ||  |||||  ||||||||||
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
              |||||||  |||||  ||||||  |  |||||  ||  ||  ||||||||      ||  ||  |||||
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
              |  ||  ||  ||||||  ||||||||  |||||  |||||  |||||||    ||  ||  ||  |  ||
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

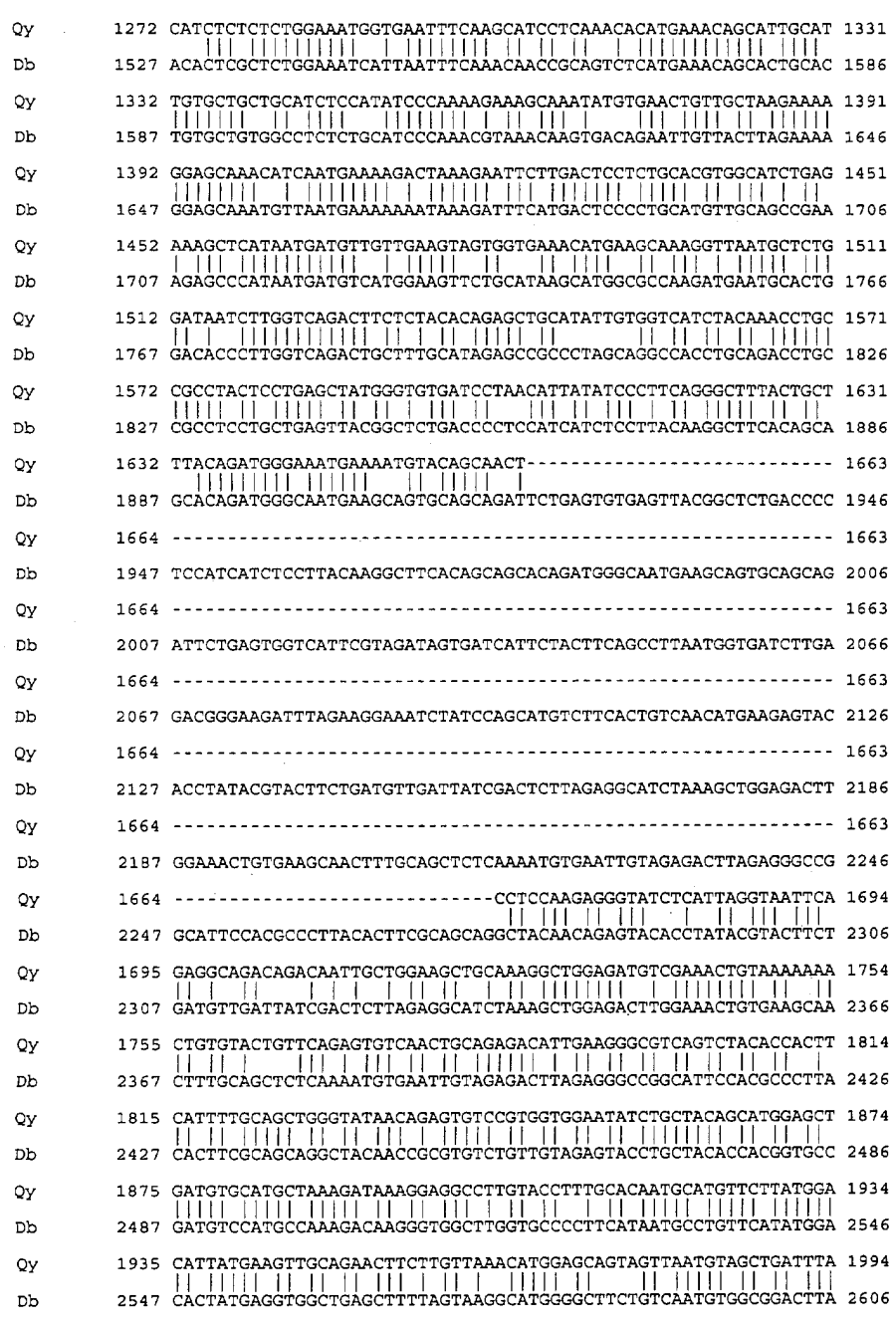
Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
              |||||  |||||  ||||||||||  ||  |||||    |  |||||||  ||  |||
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

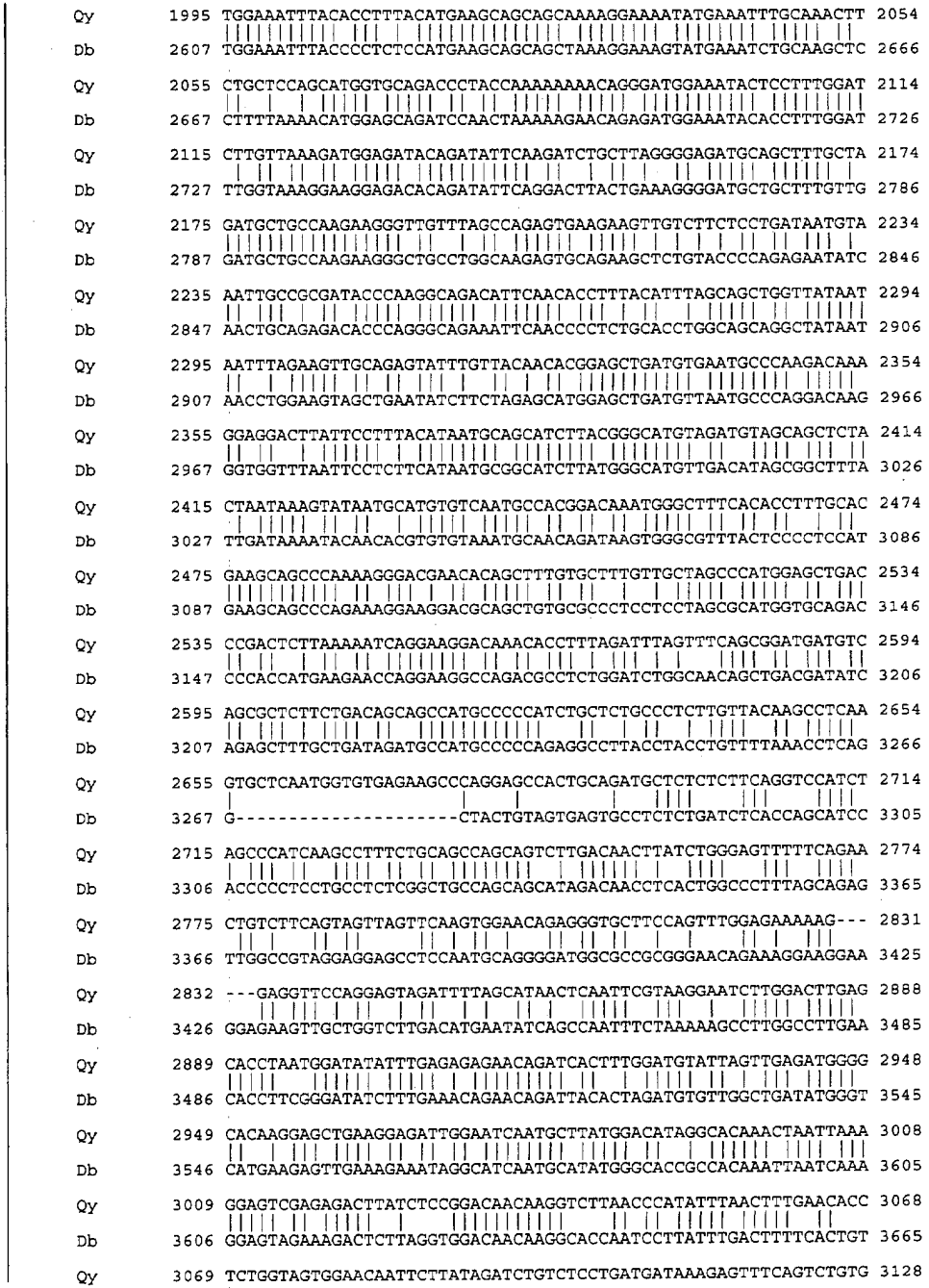
Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
              ||  |||||  |||||  |  |||||  ||  |||||  ||  ||  |||||  |||||  ||||||  |
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

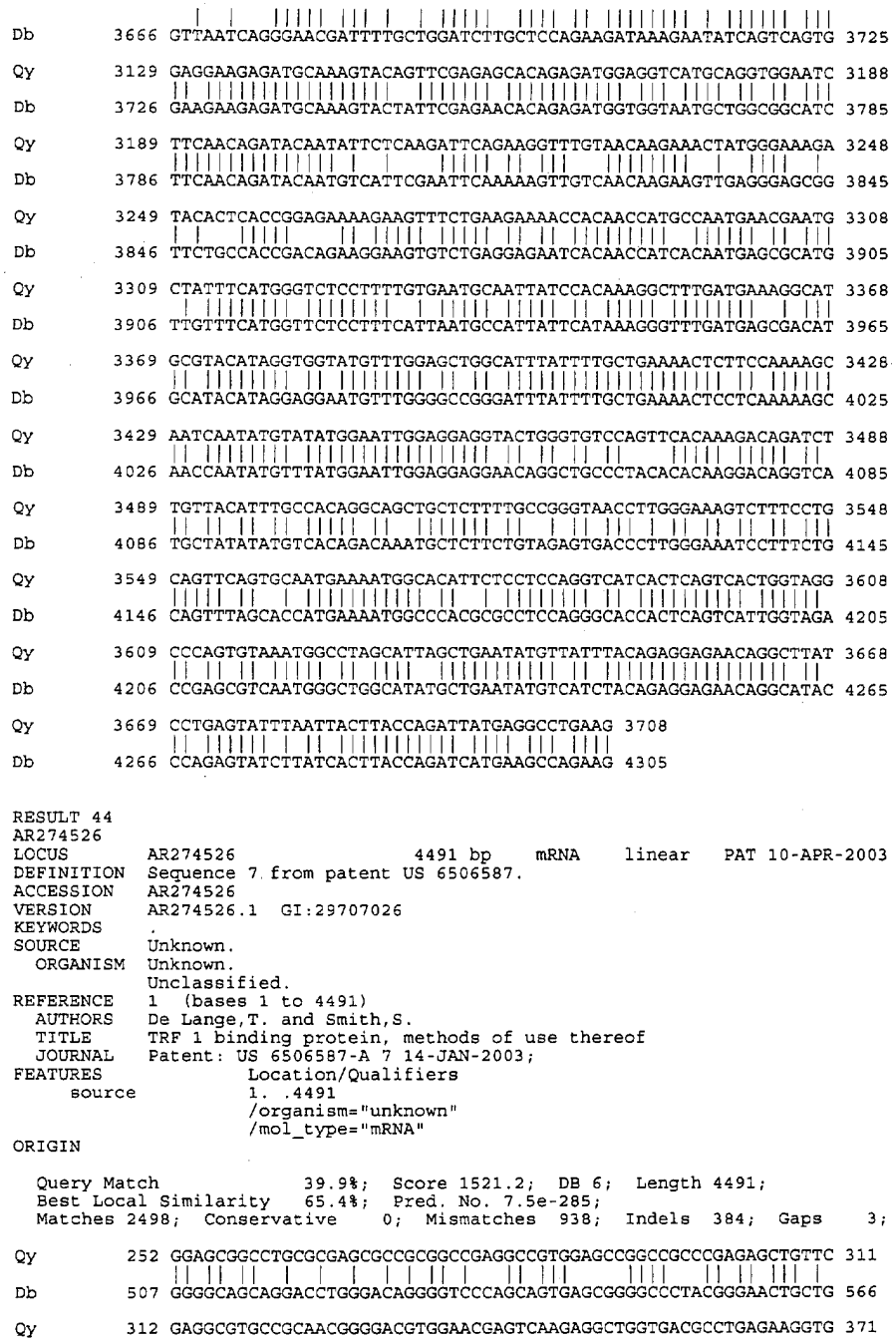
Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
              ||  |||||  ||  |||  |  |  |||  ||||  ||  ||  ||  |  ||||  ||  ||
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
                 |||||||||  |  ||  |||||||  ||  ||  ||  |  |  ||  |||||  |  ||||||
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
              |||||||  ||  ||  ||  ||||||  ||  |||||||  ||  |  ||  .  |  |  ||||||
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526
```







```
Db       3666 GTTAATCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCAGTG 3725

Qy       3129 GAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATC 3188

Db       3726 GAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGCATC 3785

Qy       3189 TTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGA 3248

Db       3786 TTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGG 3845

Qy       3249 TACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATG 3308

Db       3846 TTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGCATG 3905

Qy       3309 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 3368

Db       3906 TTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGACAT 3965

Qy       3369 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 3428

Db       3966 GCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAAAGC 4025

Qy       3429 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3488

Db       4026 AACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCA 4085

Qy       3489 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3548

Db       4086 TGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTG 4145

Qy       3549 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3608

Db       4146 CAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGA 4205

Qy       3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668

Db       4206 CCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATAC 4265

Qy       3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708

Db       4266 CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4305


RESULT 44
AR274526
LOCUS       AR274526                4491 bp    mRNA    linear   PAT 10-APR-2003
DEFINITION  Sequence 7 from patent US 6506587.
ACCESSION   AR274526
VERSION     AR274526.1  GI:29707026
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4491)
  AUTHORS   De Lange,T. and Smith,S.
  TITLE     TRF 1 binding protein, methods of use thereof
  JOURNAL   Patent: US 6506587-A 7 14-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .4491
                     /organism="unknown"
                     /mol_type="mRNA"
ORIGIN

  Query Match             39.9%;  Score 1521.2;  DB 6;  Length 4491;
  Best Local Similarity   65.4%;  Pred. No. 7.5e-285;
  Matches 2498;  Conservative    0;  Mismatches  938;  Indels  384;  Gaps    3;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311

Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371

Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431

Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491

Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551

Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611

Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671

Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731

Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791

Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851

Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911

Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971

Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031

Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091

Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151

Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211

Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271

Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

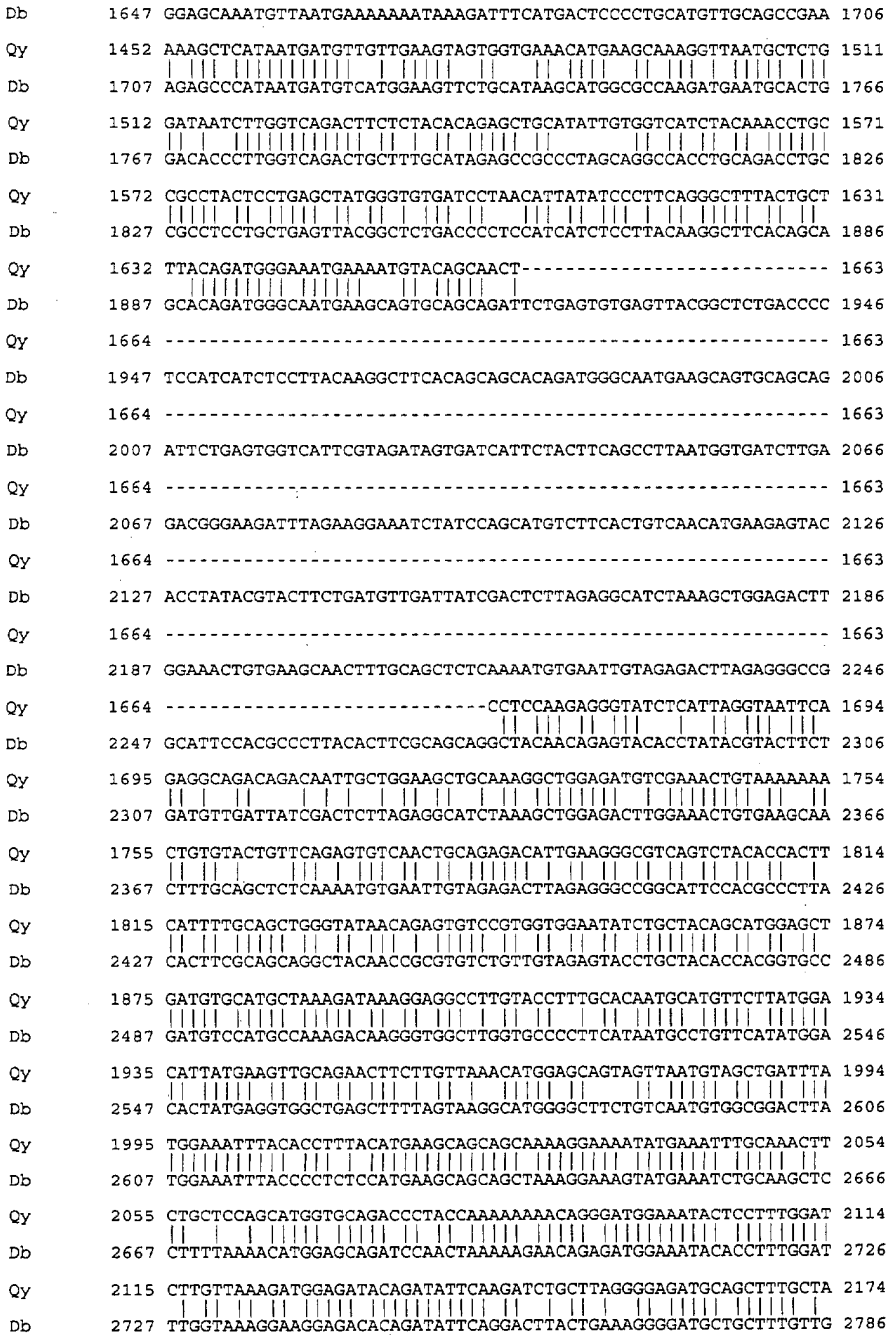
Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331

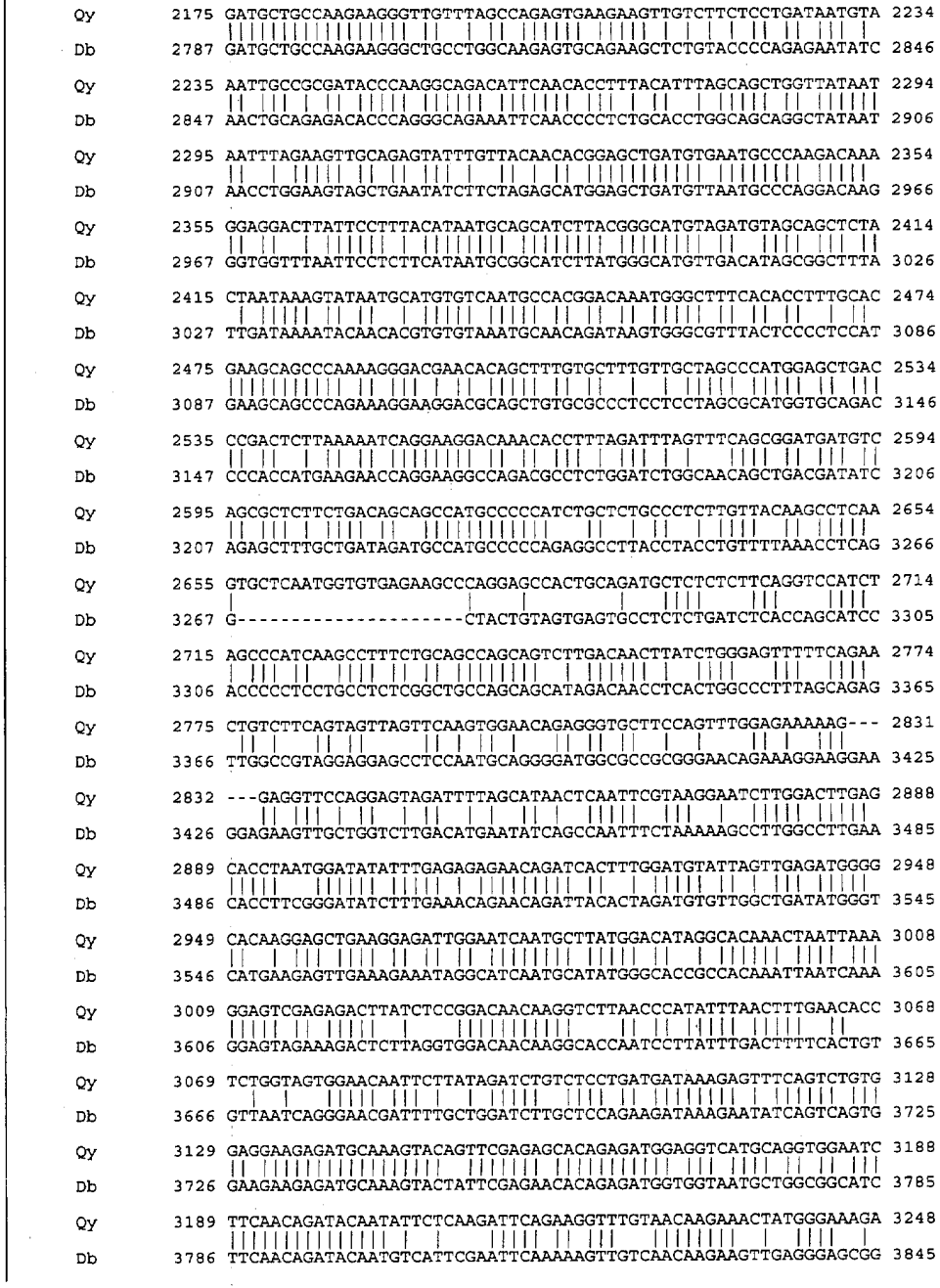
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391

Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
```





```
Qy       3249 TACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATG 3308
Db       3846 TTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGCATG 3905

Qy       3309 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 3368
Db       3906 TTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGACAT 3965

Qy       3369 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 3428
Db       3966 GCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAAAGC 4025

Qy       3429 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3488
Db       4026 AACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCA 4085

Qy       3489 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3548
Db       4086 TGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTG 4145

Qy       3549 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3608
Db       4146 CAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGA 4205

Qy       3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668
Db       4206 CCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATAC 4265

Qy       3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
Db       4266 CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4305
```

```
RESULT 45
AX467067
LOCUS       AX467067               4491 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 7 from Patent WO0229017.
ACCESSION   AX467067
VERSION     AX467067.1  GI:21900395
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   de Lange,T. and Smith,S.
  TITLE     A novel trf1 binding protein, methods of use thereof
  JOURNAL   Patent: WO 0229017-A 7 11-APR-2002;
            THE ROCKEFELLER UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1. .4491
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="cDNA"
     CDS             6. .2027
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD42445.1"
                     /db_xref="GI:21900396"
                     /db_xref="REMTREMBL:CAD42445"
                     /translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPPLSPGLAPGTT
                     PASPTASGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGTSCCSTTSTICTVAAA
                     PVVPAVSTSSAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPGSSLAESPEAAGVS
                     STAPLGPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSS
                     PLHFAAGFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPN
                     ARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEY
                     KKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHG
                     ADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEV
                     CSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKT
                     LALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
                     ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQG
                     FTAAQMGNEAVQQILSVSYGSDPSIISLQGFTAAQMGNEAVQQILSGHS"
```

```
ORIGIN

  Query Match             39.9%;  Score 1521.2;  DB 6;  Length 4491;
  Best Local Similarity   65.4%;  Pred. No. 7.5e-285;
  Matches 2498;  Conservative    0;  Mismatches  938;  Indels  384;  Gaps    3;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

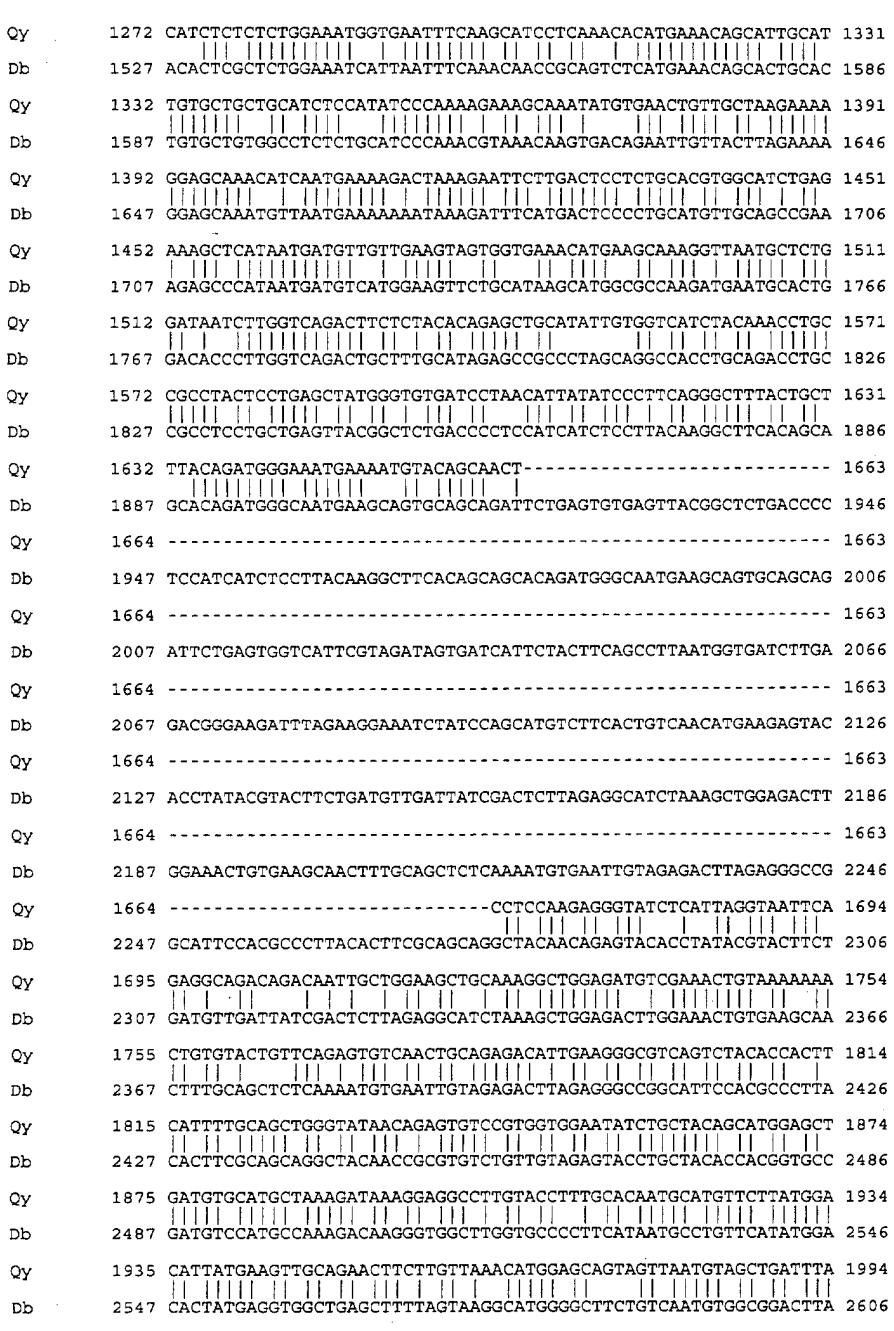
Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

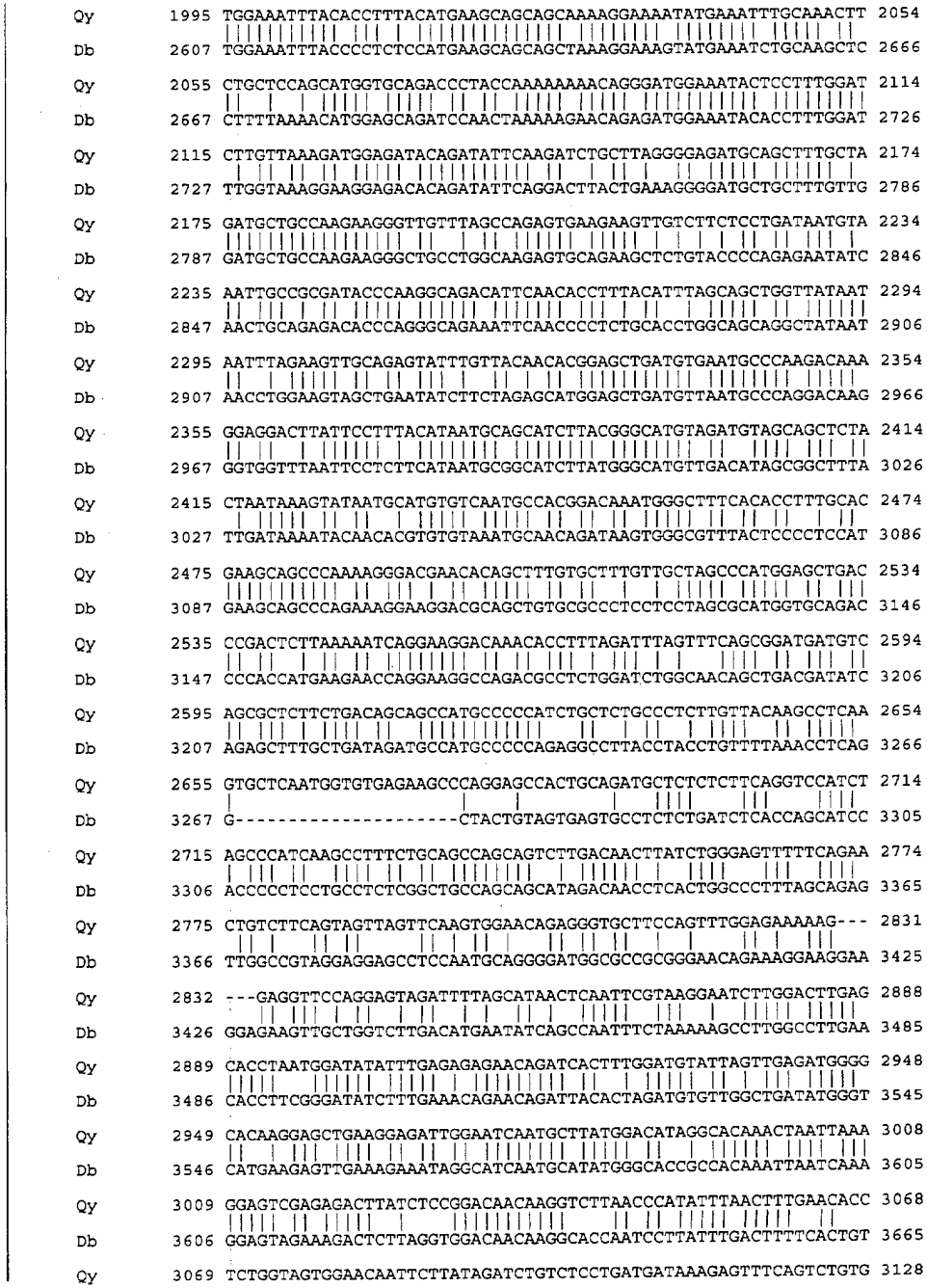
Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526
```





```
Db       3666 GTTAATCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCAGTG 3725

Qy       3129 GAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATC 3188

Db       3726 GAAGAAGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGCATC 3785

Qy       3189 TTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGA 3248

Db       3786 TTCAACAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGG 3845

Qy       3249 TACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATG 3308

Db       3846 TTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGCATG 3905

Qy       3309 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 3368

Db       3906 TTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGACAT 3965

Qy       3369 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 3428

Db       3966 GCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAAAGC 4025

Qy       3429 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3488

Db       4026 AACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCA 4085

Qy       3489 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3548

Db       4086 TGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTG 4145

Qy       3549 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3608

Db       4146 CAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGA 4205

Qy       3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668

Db       4206 CCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATAC 4265

Qy       3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708

Db       4266 CCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4305


Search completed: March 30, 2004, 09:40:36
Job time : 14465.4 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 01:40:08 ; Search time 1288.71 Seconds
(without alignments)
11024.221 Million cell updates/sec

Title: US-09-843-159B-2
Perfect score: 3816
Sequence: 1 cgcgctgctccgcccgccgc.....cctttgctgaaaaaaaaaa 3816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3791.4	99.4	4275	10	US-09-972-115A-5	Sequence 5, Appli
2	3791.4	99.4	4406	14	US-10-199-937-106	Sequence 106, App
3	3791.4	99.4	4992	14	US-10-199-937-134	Sequence 134, App
4	3791.4	99.4	5002	14	US-10-199-937-132	Sequence 132, App
5	3788.4	99.3	6018	10	US-09-849-602-11	Sequence 11, Appl
6	3770.8	98.8	5075	12	US-10-302-172-556	Sequence 556, App
7	3595.8	94.2	4127	14	US-10-199-937-100	Sequence 100, App
8	3587.4	94.0	5778	14	US-10-163-587A-14	Sequence 14, Appl
9	3504.8	91.8	3508	14	US-10-199-937-1	Sequence 1, Appli
10	3341	87.6	4297	10	US-09-972-115A-3	Sequence 3, Appli
11	3297	86.4	3400	9	US-09-509-196A-1	Sequence 1, Appli
12	3039	79.6	4493	10	US-09-972-115A-1	Sequence 1, Appli
13	2439.8	63.9	2971	14	US-10-199-937-92	Sequence 92, Appl
14	2439.8	63.9	3353	14	US-10-199-937-88	Sequence 88, Appl
15	2357.8	61.8	3799	14	US-10-199-937-90	Sequence 90, Appl
16	1885	49.4	3984	14	US-10-199-937-3	Sequence 3, Appli
17	1885	49.4	4134	9	US-09-841-835-1	Sequence 1, Appli
18	1885	49.4	4134	10	US-09-972-115A-7	Sequence 7, Appli
19	1568	41.1	1568	14	US-10-199-937-98	Sequence 98, Appl
20	1533.8	40.2	4435	14	US-10-252-157-342	Sequence 342, App
21	1521.2	39.9	4491	9	US-09-841-835-7	Sequence 7, Appli
22	1369.4	35.9	4657	9	US-09-841-835-9	Sequence 9, Appli
23	1240.4	32.5	1466	12	US-10-276-774-340	Sequence 340, App
24	1192.6	31.3	2409	9	US-09-964-899-40	Sequence 40, Appl
25	1168.4	30.6	3308	14	US-10-199-937-177	Sequence 177, App
26	1084.8	28.4	1103	14	US-10-199-937-159	Sequence 159, App
27	1066.8	28.0	5482	14	US-10-199-937-138	Sequence 138, App
28	794.4	20.8	797	14	US-10-252-157-341	Sequence 341, App
29	574	15.0	582	14	US-10-199-937-45	Sequence 45, Appl
30	563.6	14.8	647	14	US-10-199-937-104	Sequence 104, App
c 31	518.4	13.6	520	10	US-09-972-115A-10	Sequence 10, Appl
32	467.6	12.3	1069	9	US-09-833-381-841	Sequence 841, App
33	461	12.1	465	9	US-09-833-381-1153	Sequence 1153, Ap
34	459	12.0	460	14	US-10-106-698-2223	Sequence 2223, Ap
35	459	12.0	460	14	US-10-199-937-5	Sequence 5, Appli
36	396.8	10.4	402	14	US-10-181-447A-11	Sequence 11, Appl
37	380.4	10.0	384	10	US-09-972-115A-9	Sequence 9, Appli
38	368	9.6	441	10	US-09-918-995-17041	Sequence 17041, A
39	359.6	9.4	1691	14	US-10-199-937-41	Sequence 41, Appl
40	359.6	9.4	1692	14	US-10-199-937-43	Sequence 43, Appl
41	356	9.3	356	14	US-10-199-937-49	Sequence 49, Appl
42	354.6	9.3	564	14	US-10-199-937-7	Sequence 7, Appli
43	346	9.1	392	14	US-10-199-937-14	Sequence 14, Appl
c 44	340.4	8.9	353	14	US-10-199-937-12	Sequence 12, Appl
45	323	8.5	353	13	US-10-040-739-1179	Sequence 1179, Ap

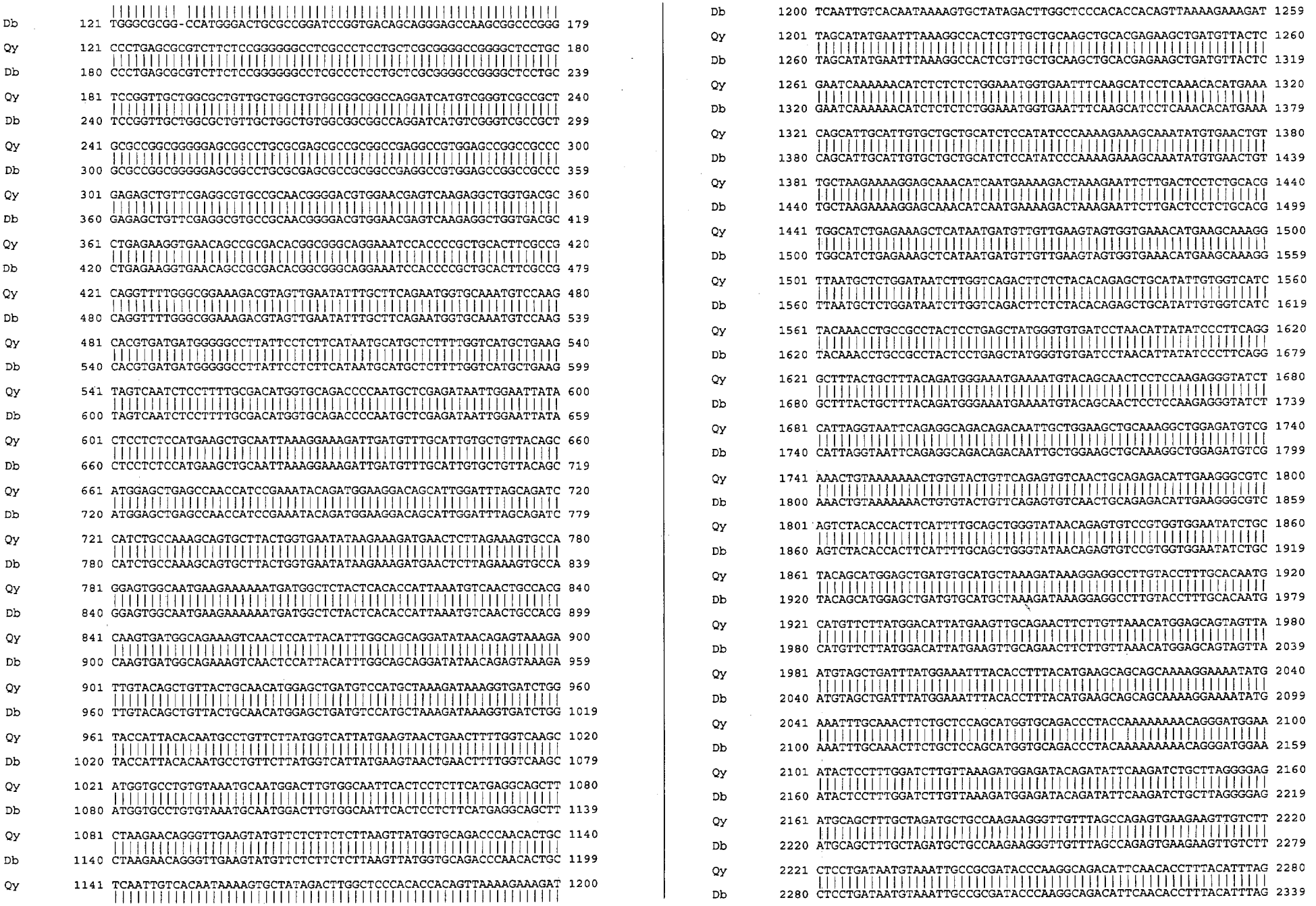
ALIGNMENTS

RESULT 1
US-09-972-115A-5
; Sequence 5, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 4275
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (284)..(3781)
; OTHER INFORMATION:
US-09-972-115A-5

Query Match 99.4%; Score 3791.4; DB 10; Length 4275;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAGGGGCGCGCG 120

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
```


Db 121 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 179
Qy 121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
Db 180 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 239
Qy 181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
Db 240 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 299
Qy 241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db 300 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 359
Qy 301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db 360 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 419
Qy 361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db 420 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 479
Qy 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db 480 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 539
Qy 481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db 540 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 599
Qy 541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db 600 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 659
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db 660 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 719
Qy 661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db 720 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 779
Qy 721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db 780 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 839
Qy 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db 840 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 899
Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 900 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 959
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 960 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1019
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1020 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1079
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1080 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1139
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1140 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1199
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1200 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1259
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1260 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1319
Qy 1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1320 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1379
Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1380 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1439
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 1440 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1499
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 1500 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1559
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 1560 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1619
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 1620 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1679
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 1680 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1739
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 1740 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1799
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1800 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1859
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 1860 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1919
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 1920 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1979
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 1980 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2039
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2040 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2099
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db 2100 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2159
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2160 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2219
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2220 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2279
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2280 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2339

```
Qy       2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2340 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2399

Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2400 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2459

Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2460 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2519

Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2520 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2579

Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2580 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2639

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2640 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2699

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2700 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2759

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2760 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2819

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2820 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2879

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2880 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2939

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2940 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2999

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3000 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3059

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3060 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3119

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3120 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3179

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3180 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3239

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3240 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3299

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3300 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3359

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3360 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3419

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3420 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3479

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3480 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3539

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3540 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3599

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3600 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3659

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3660 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3719

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3720 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3779

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3780 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3839

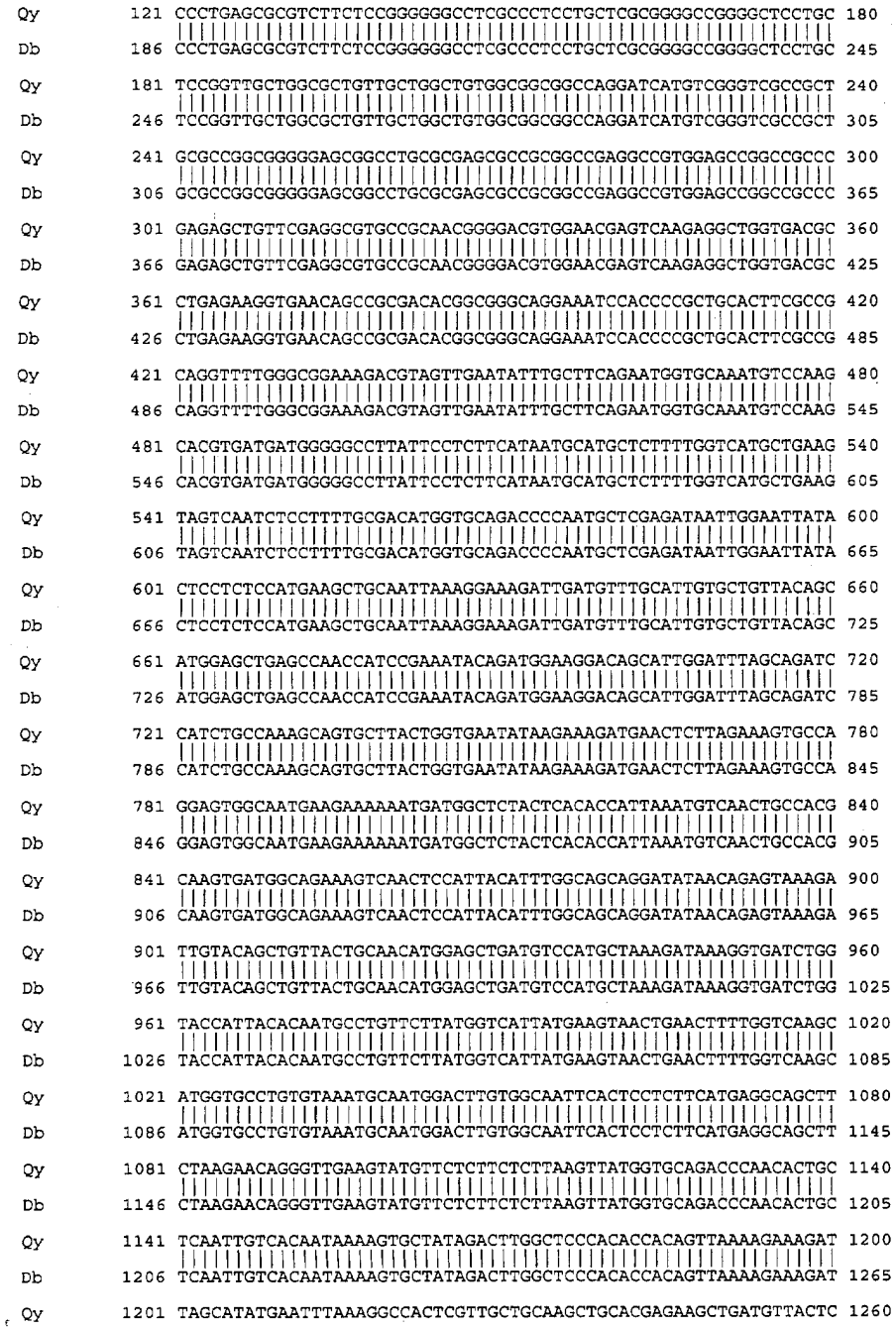
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||
Db       3840 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 2
US-10-199-937-106
; Sequence 106, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 4406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(3787)
US-10-199-937-106

Query Match 99.4%; Score 3791.4; DB 14; Length 4406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              |||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 126

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 185
```


Qy 121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
Db 186 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 245
Qy 181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
Db 246 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 305
Qy 241 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db 306 GCGCCGGCGGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 365
Qy 301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db 366 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 425
Qy 361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db 426 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 485
Qy 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db 486 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 545
Qy 481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db 546 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 605
Qy 541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db 606 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 665
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db 666 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 725
Qy 661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db 726 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 785
Qy 721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db 786 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 845
Qy 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db 846 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 905
Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 906 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 965
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 966 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1025
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1026 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1085
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1086 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1145
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1146 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1205
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1206 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1265
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1266 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1325
Qy 1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1326 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1385
Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1386 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1445
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 1446 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1505
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 1506 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1565
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 1566 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1625
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 1626 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1685
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 1686 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1745
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 1746 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1805
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 1806 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1865
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 1866 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1925
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 1926 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1985
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 1986 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2045
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2046 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2105
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db 2106 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2165
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2166 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2225
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2226 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2285
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2286 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2345
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340

```
Db       2346 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2405
Qy       2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db       2406 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2465
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       2466 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2525
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       2526 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2585
Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       2586 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2645
Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       2646 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2705
Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       2706 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2765
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       2766 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2825
Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       2826 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2885
Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       2886 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2945
Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       2946 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3005
Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       3006 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3065
Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3066 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3125
Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3126 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3185
Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3186 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3245
Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3246 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3305
Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3306 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3365
Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3366 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3425
Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       3426 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3485

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       3486 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3545
Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       3546 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3605
Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       3606 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3665
Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       3666 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3725
Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       3726 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3785
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       3786 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3845
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3846 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3878
```

RESULT 3
US-10-199-937-134
; Sequence 134, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
; LENGTH: 4992
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (876)..(4373)
US-10-199-937-134

Query Match 99.4%; Score 3791.4; DB 14; Length 4992;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
Db        653 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 712
Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
Db        713 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 771
Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
Db        772 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 831
```

```
Qy       181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
Db       832 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 891

Qy       241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db       892 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 951

Qy       301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db       952 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 1011

Qy       361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db      1012 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 1071

Qy       421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db      1072 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 1131

Qy       481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db      1132 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 1191

Qy       541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db      1192 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 1251

Qy       601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db      1252 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 1311

Qy       661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db      1312 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 1371

Qy       721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db      1372 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 1431

Qy       781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db      1432 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 1491

Qy       841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db      1492 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1551

Qy       901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db      1552 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1611

Qy       961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db      1612 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1671

Qy      1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db      1672 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1731

Qy      1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db      1732 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1791

Qy      1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db      1792 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1851

Qy      1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db      1852 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1911

Qy      1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db      1912 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1971

Qy      1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db      1972 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 2031

Qy      1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db      2032 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 2091

Qy      1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db      2092 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 2151

Qy      1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db      2152 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 2211

Qy      1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db      2212 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 2271

Qy      1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db      2272 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 2331

Qy      1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db      2332 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 2391

Qy      1741 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db      2392 AAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 2451

Qy      1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db      2452 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2511

Qy      1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db      2512 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2571

Qy      1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db      2572 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2631

Qy      1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db      2632 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2691

Qy      2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db      2692 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2751

Qy      2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db      2752 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2811

Qy      2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db      2812 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2871

Qy      2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db      2872 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2931

Qy      2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db      2932 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2991

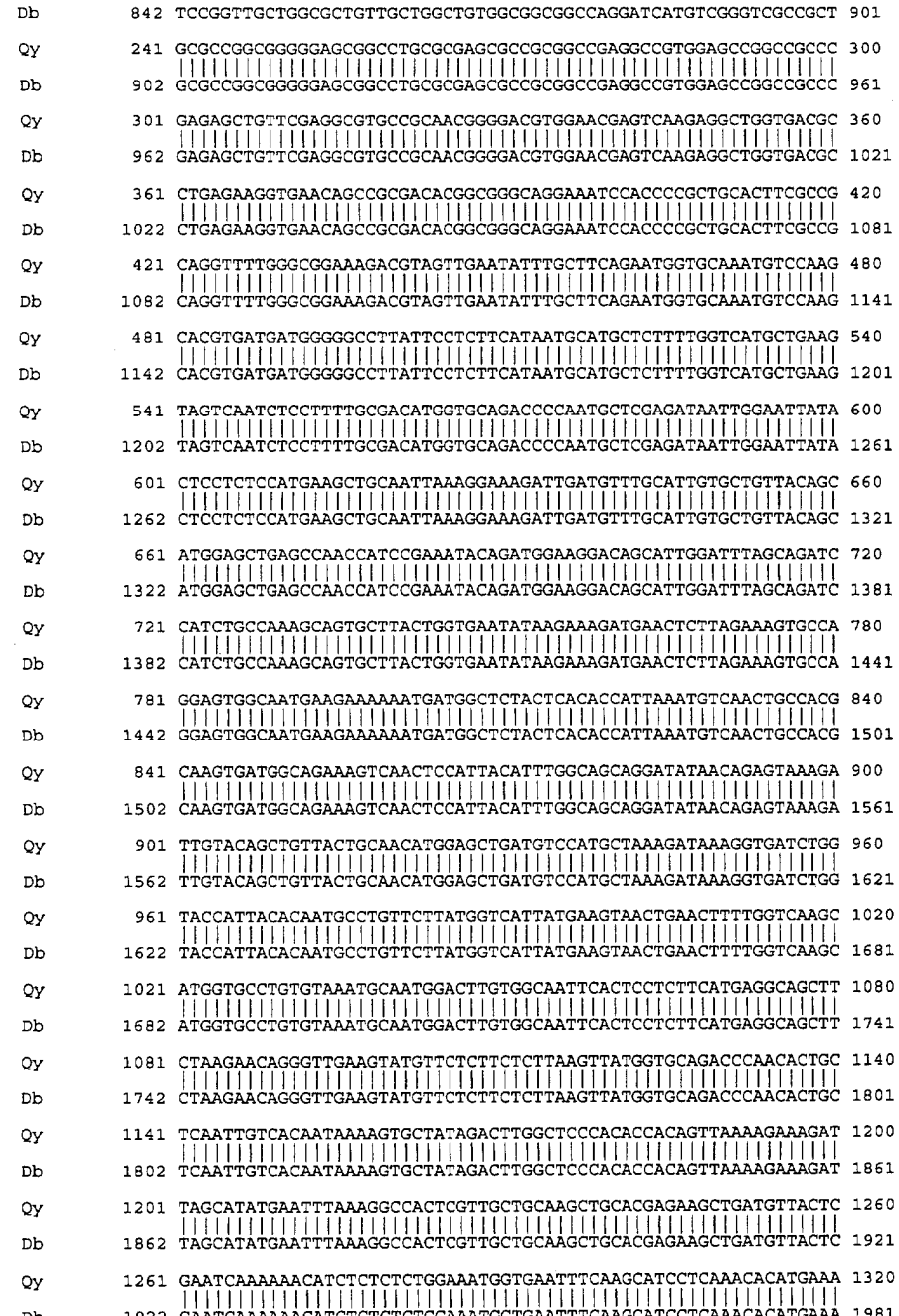
Qy      2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
```

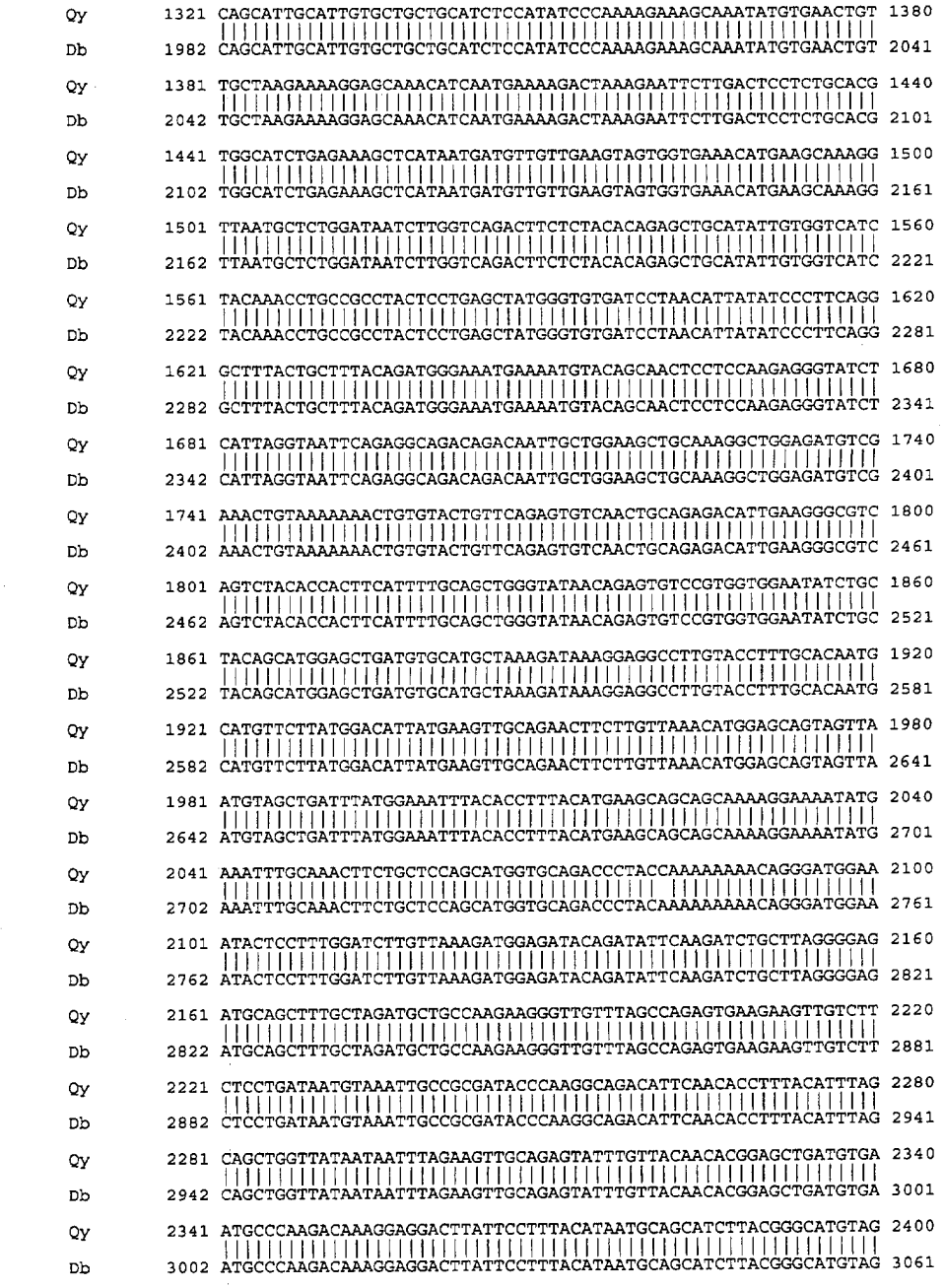
```
Db       2992 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3051
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db       3052 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 3111
Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db       3112 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 3171
Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db       3172 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3231
Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db       3232 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3291
Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db       3292 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 3351
Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db       3352 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 3411
Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db       3412 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 3471
Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db       3472 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3531
Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db       3532 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3591
Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db       3592 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3651
Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
Db       3652 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3711
Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
Db       3712 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3771
Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
Db       3772 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3831
Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
Db       3832 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3891
Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
Db       3892 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3951
Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
Db       3952 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 4011
Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
Db       4012 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 4071
Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
Db       4072 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 4131
Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
Db       4132 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 4191
Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
Db       4192 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 4251
Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
Db       4252 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 4311
Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
Db       4312 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 4371
Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
Db       4372 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 4431
Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       4432 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4464
```

RESULT 4
US-10-199-937-132
; Sequence 132, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 5002
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)..(4383)
US-10-199-937-132

Query Match 99.4%; Score 3791.4; DB 14; Length 5002;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3806; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
Db        663 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 722
Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
Db        723 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 781
Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
Db        782 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 841
Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
```


Db 842 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 901
Qy 241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
Db 902 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 961
Qy 301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
Db 962 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 1021
Qy 361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
Db 1022 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 1081
Qy 421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
Db 1082 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 1141
Qy 481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
Db 1142 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 1201
Qy 541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
Db 1202 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 1261
Qy 601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
Db 1262 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 1321
Qy 661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
Db 1322 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 1381
Qy 721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
Db 1382 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 1441
Qy 781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
Db 1442 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 1501
Qy 841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
Db 1502 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 1561
Qy 901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
Db 1562 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1621
Qy 961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
Db 1622 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1681
Qy 1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
Db 1682 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1741
Qy 1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
Db 1742 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1801
Qy 1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
Db 1802 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1861
Qy 1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
Db 1862 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1921
Qy 1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
Db 1922 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1981


Qy 1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
Db 1982 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 2041
Qy 1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
Db 2042 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 2101
Qy 1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
Db 2102 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 2161
Qy 1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
Db 2162 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 2221
Qy 1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620
Db 2222 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 2281
Qy 1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680
Db 2282 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 2341
Qy 1681 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1740
Db 2342 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 2401
Qy 1741 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1800
Db 2402 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 2461
Qy 1801 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1860
Db 2462 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 2521
Qy 1861 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1920
Db 2522 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 2581
Qy 1921 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1980
Db 2582 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2641
Qy 1981 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2040
Db 2642 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2701
Qy 2041 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAA 2100
Db 2702 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2761
Qy 2101 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2160
Db 2762 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2821
Qy 2161 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2220
Db 2822 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2881
Qy 2221 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2280
Db 2882 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2941
Qy 2281 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2340
Db 2942 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 3001
Qy 2341 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2400
Db 3002 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 3061

```
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 3121

Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 3181

Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 3241

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 CAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 3301

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 3361

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 3421

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 3481

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 3541

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3601

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3661

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3721

Qy       3061 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3781

Qy       3121 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3782 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3841

Qy       3181 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3842 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3901

Qy       3241 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3902 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3961

Qy       3301 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3962 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 4021

Qy       3361 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4022 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 4081

Qy       3421 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4082 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 4141

Qy       3481 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4142 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 4201

Qy       3541 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4202 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 4261

Qy       3601 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4262 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 4321

Qy       3661 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4322 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 4381

Qy       3721 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4382 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 4441

Qy       3781 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||
Db       4442 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4474
```

RESULT 5
US-09-849-602-11
; Sequence 11, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 6018
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-602-11

Query Match 99.3%; Score 3788.4; DB 10; Length 6018;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3803; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```
Qy          4 GCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCGTGG 63
              ||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCGTGG 60

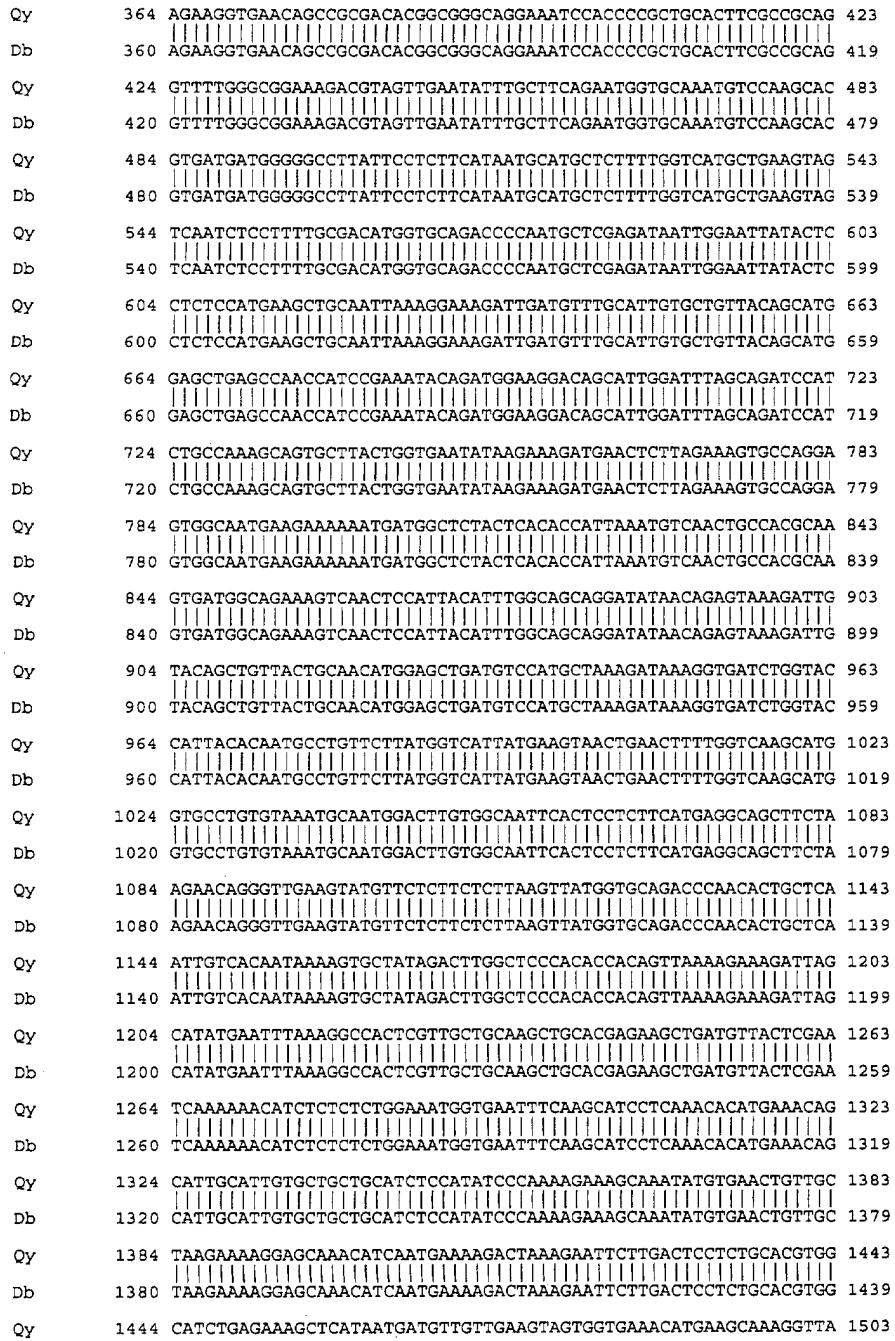
Qy         64 GCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCC 123
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCC 119

Qy        124 TGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 TGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCC 179

Qy        184 GGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCG 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 GGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCG 239

Qy        244 CCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAG 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAG 299

Qy        304 AGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTG 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 AGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTG 359
```


Qy 364 AGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAG 423
Db 360 AGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAG 419
Qy 424 GTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCAC 483
Db 420 GTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCAC 479
Qy 484 GTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAG 543
Db 480 GTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAG 539
Qy 544 TCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTC 603
Db 540 TCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTC 599
Qy 604 CTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATG 663
Db 600 CTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATG 659
Qy 664 GAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCAT 723
Db 660 GAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCAT 719
Qy 724 CTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGA 783
Db 720 CTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGA 779
Qy 784 GTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAA 843
Db 780 GTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAA 839
Qy 844 GTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTG 903
Db 840 GTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTG 899
Qy 904 TACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTAC 963
Db 900 TACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTAC 959
Qy 964 CATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATG 1023
Db 960 CATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATG 1019
Qy 1024 GTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTA 1083
Db 1020 GTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTA 1079
Qy 1084 AGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCA 1143
Db 1080 AGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCA 1139
Qy 1144 ATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAG 1203
Db 1140 ATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAG 1199
Qy 1204 CATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAA 1263
Db 1200 CATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAA 1259
Qy 1264 TCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAG 1323
Db 1260 TCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAG 1319
Qy 1324 CATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGC 1383
Db 1320 CATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGC 1379
Qy 1384 TAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGG 1443
Db 1380 TAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGG 1439
Qy 1444 CATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTA 1503
Db 1440 CATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTA 1499
Qy 1504 ATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTAC 1563
Db 1500 ATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTAC 1559
Qy 1564 AAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCT 1623
Db 1560 AAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCT 1619
Qy 1624 TTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCAT 1683
Db 1620 TTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCAT 1679
Qy 1684 TAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAA 1743
Db 1680 TAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAA 1739
Qy 1744 CTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGT 1803
Db 1740 CTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGT 1799
Qy 1804 CTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTAC 1863
Db 1800 CTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTAC 1859
Qy 1864 AGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCAT 1923
Db 1860 AGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCAT 1919
Qy 1924 GTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATG 1983
Db 1920 GTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATG 1979
Qy 1984 TAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAA 2043
Db 1980 TAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAA 2039
Qy 2044 TTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATA 2103
Db 2040 TTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATA 2099
Qy 2104 CTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATG 2163
Db 2100 CTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATG 2159
Qy 2164 CAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTC 2223
Db 2160 CAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTC 2219
Qy 2224 CTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAG 2283
Db 2220 CTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAG 2279
Qy 2284 CTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATG 2343
Db 2280 CTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATG 2339
Qy 2344 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATG 2403
Db 2340 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATG 2399
Qy 2404 TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCA 2463
Db 2400 TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCA 2459
Qy 2464 CACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCC 2523
Db 2460 CACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCC 2519
Qy 2524 ATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAG 2583

```
Db       2520 ATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAG 2579

Qy       2584 CGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTT 2643

Db       2580 CAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTT 2639

Qy       2644 ACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTT 2703

Db       2640 ACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTT 2699

Qy       2704 CAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGA 2763

Db       2700 CAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGA 2759

Qy       2764 GTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGG 2823

Db       2760 GTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGG 2819

Qy       2824 AGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGAC 2883

Db       2820 AGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGAC 2879

Qy       2884 TTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGA 2943

Db       2880 TTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGA 2939

Qy       2944 TGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAA 3003

Db       2940 TGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAA 2999

Qy       3004 TTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGA 3063

Db       3000 TTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGA 3059

Qy       3064 ACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGT 3123

Db       3060 ACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGT 3119

Qy       3124 CTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTG 3183

Db       3120 CTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTG 3179

Qy       3184 GAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGG 3243

Db       3180 GAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGG 3239

Qy       3244 AAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAAC 3303

Db       3240 AAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAAC 3299

Qy       3304 GAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAA 3363

Db       3300 GAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAA 3359

Qy       3364 GGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCA 3423

Db       3360 GGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCA 3419

Qy       3424 AAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACA 3483

Db       3420 AAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACA 3479

Qy       3484 GATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTT 3543

Db       3480 GATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTT 3539

Qy       3544 TCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTG 3603

Db       3540 TCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTG 3599

Qy       3604 GTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGG 3663

Db       3600 GTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGG 3659

Qy       3664 CTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGAT 3723

Db       3660 CTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGAT 3719

Qy       3724 AAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCC 3783

Db       3720 AAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCC 3779

Qy       3784 TCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813

Db       3780 TCTACGTTTTACTCCTTTGCTGAAAAAAAA 3809


RESULT 6
US-10-302-172-556
; Sequence 556, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 803_1CNCP
;  CURRENT APPLICATION NUMBER: US/10/302,172
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/225,251
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: PCT US02/05095
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 09/799,451
;  PRIOR FILING DATE: 2001-03-05
;  NUMBER OF SEQ ID NOS: 950
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 556
;   LENGTH: 5075
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (253)..(3750)
US-10-302-172-556

  Query Match             98.8%;  Score 3770.8;  DB 12;  Length 5075;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 3800;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    2;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60

Db         29 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 88

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120

Db         89 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 147

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCT-CGCGGGGCCCGGGGCTCCTG 179

Db        148 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCCCCGGGGCCGGGGCTCCTG 207

Qy        180 CTCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGC 239

Db        208 CTCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGC 267

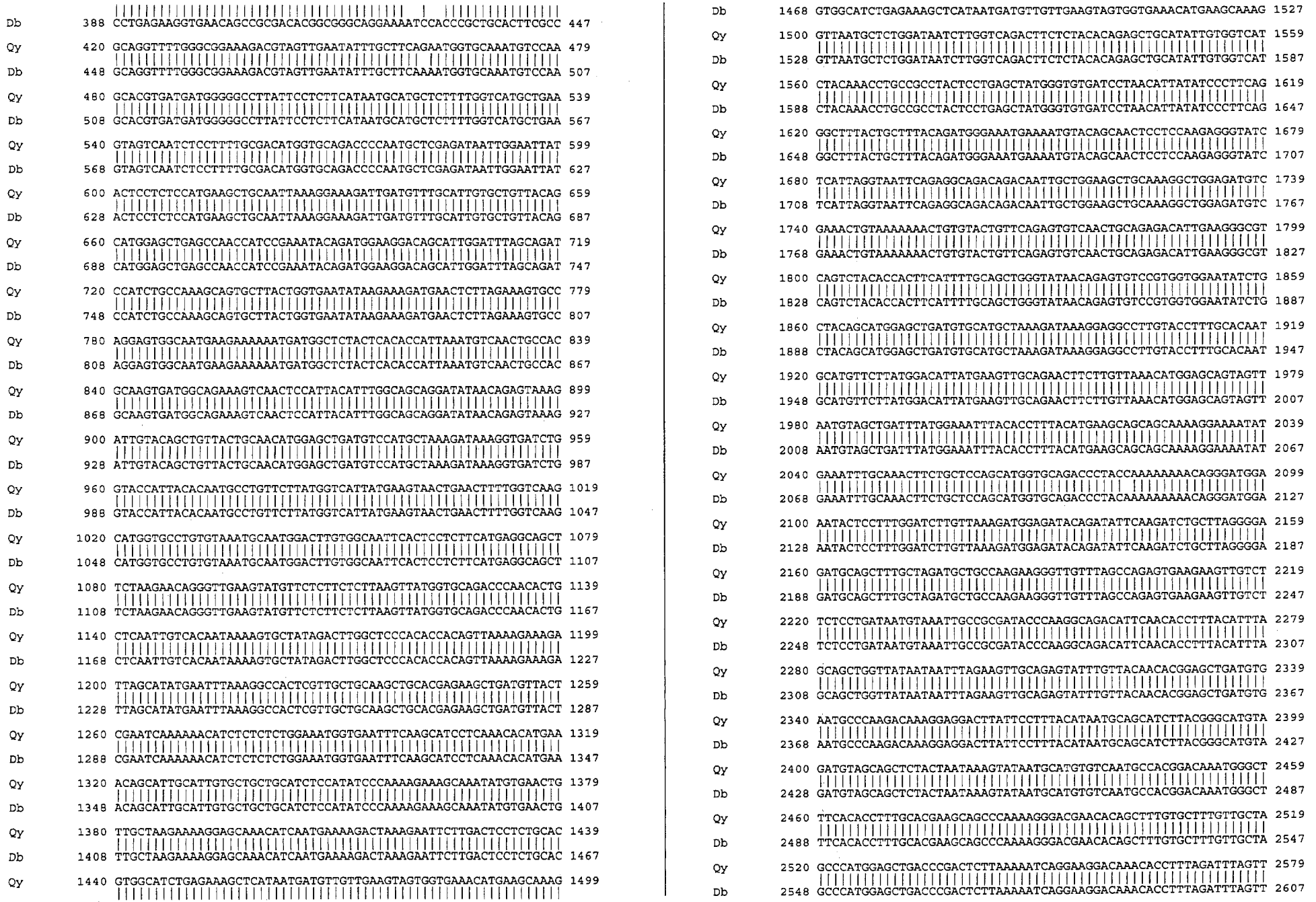
Qy        240 TGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCC 299

Db        268 TGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCC 327

Qy        300 CGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACG 359

Db        328 CGAGAGCTGTTCGAGGCGTGCCCCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACG 387

Qy        360 CCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCC 419
```



```
Qy       2580 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2608 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2667

Qy       2640 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2699
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2668 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2727

Qy       2700 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2759
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2728 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2787

Qy       2760 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2819
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2788 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2847

Qy       2820 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2879
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2848 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2907

Qy       2880 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2939
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2908 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2967

Qy       2940 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 2999
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2968 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 3027

Qy       3000 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3059
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3028 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3087

Qy       3060 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3088 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3147

Qy       3120 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3148 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3207

Qy       3180 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3267

Qy       3240 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3268 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3327

Qy       3300 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3328 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3387

Qy       3360 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3388 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3447

Qy       3420 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3507

Qy       3480 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3567

Qy       3540 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3568 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3627

Qy       3600 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAA 3659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3628 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAA 3687

Qy       3660 CAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGAT 3719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3688 CAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGAT 3747

Qy       3720 GGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGT 3779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3748 GGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGT 3807

Qy       3780 GGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||
Db       3808 GGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3841
```

RESULT 7
US-10-199-937-100
; Sequence 100, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 4127
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(3508)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (3509)..(4127)
US-10-199-937-100

Query Match 94.2%; Score 3595.8; DB 14; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3597; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
```

```
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420
Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480
Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540
Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600
Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660
Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720
Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780
Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840
Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900
Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
Db        901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960
Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
Db        961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020
Qy       1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
Db       1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080
Qy       1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
Db       1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140
Qy       1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
Db       1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200
Qy       1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
Db       1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260
Qy       1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
Db       1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320
Qy       1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
Db       1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380
Qy       1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
Db       1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440
Qy       1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
```

```
Db       1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500
Qy       1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
Db       1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560
Qy       1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1834
Db       1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1620
Qy       1835 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894
Db       1621 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680
Qy       1895 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1954
Db       1681 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1740
Qy       1955 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 2014
Db       1741 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 1800
Qy       2015 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
Db       1801 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860
Qy       2075 CCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 2134
Db       1861 CCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 1920
Qy       2135 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
Db       1921 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980
Qy       2195 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2254
Db       1981 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2040
Qy       2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2314
Db       2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2100
Qy       2315 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
Db       2101 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160
Qy       2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2434
Db       2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2220
Qy       2435 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
Db       2221 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280
Qy       2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
Db       2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340
Qy       2555 AGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
Db       2341 AGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGC 2400
Qy       2615 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2674
Db       2401 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2460
Qy       2675 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2734
Db       2461 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2520
Qy       2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
Db       2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580
```

```
Qy       2795 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2854
Db       2581 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2640

Qy       2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2914
Db       2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2700

Qy       2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2974
Db       2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2760

Qy       2975 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 3034
Db       2761 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 2820

Qy       3035 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 3094
Db       2821 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 2880

Qy       3095 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 3154
Db       2881 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 2940

Qy       3155 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3214
Db       2941 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3000

Qy       3215 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3274
Db       3001 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3060

Qy       3275 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3334
Db       3061 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3120

Qy       3335 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3394
Db       3121 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3180

Qy       3395 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3454
Db       3181 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3240

Qy       3455 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3514
Db       3241 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3300

Qy       3515 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3574
Db       3301 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3360

Qy       3575 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3634
Db       3361 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3420

Qy       3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3694
Db       3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3480

Qy       3695 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3754
Db       3481 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3540

Qy       3755 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3541 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3599

RESULT 8
US-10-163-587A-14
; Sequence 14, Application US/10163587A
; Publication No. US20030096263A1
; GENERAL INFORMATION:
;  APPLICANT: Oliveira, Marcos
;  TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZING A
;  FILE REFERENCE: 50229-306
;  CURRENT APPLICATION NUMBER: US/10/163,587A
;  CURRENT FILING DATE:  2003-01-10
;  PRIOR APPLICATION NUMBER: 60/296,110
;  PRIOR FILING DATE: 2001-06-07
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 5778
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3501)
;   OTHER INFORMATION:
US-10-163-587A-14

  Query Match           94.0%;  Score 3587.4;  DB 14;  Length 5778;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3588;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
Db          1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60

Qy        285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
Db         61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120

Qy        345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
Db        121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180

Qy        405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
Db        181 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240

Qy        465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
Db        241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300

Qy        525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
Db        301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360

Qy        585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
Db        361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420

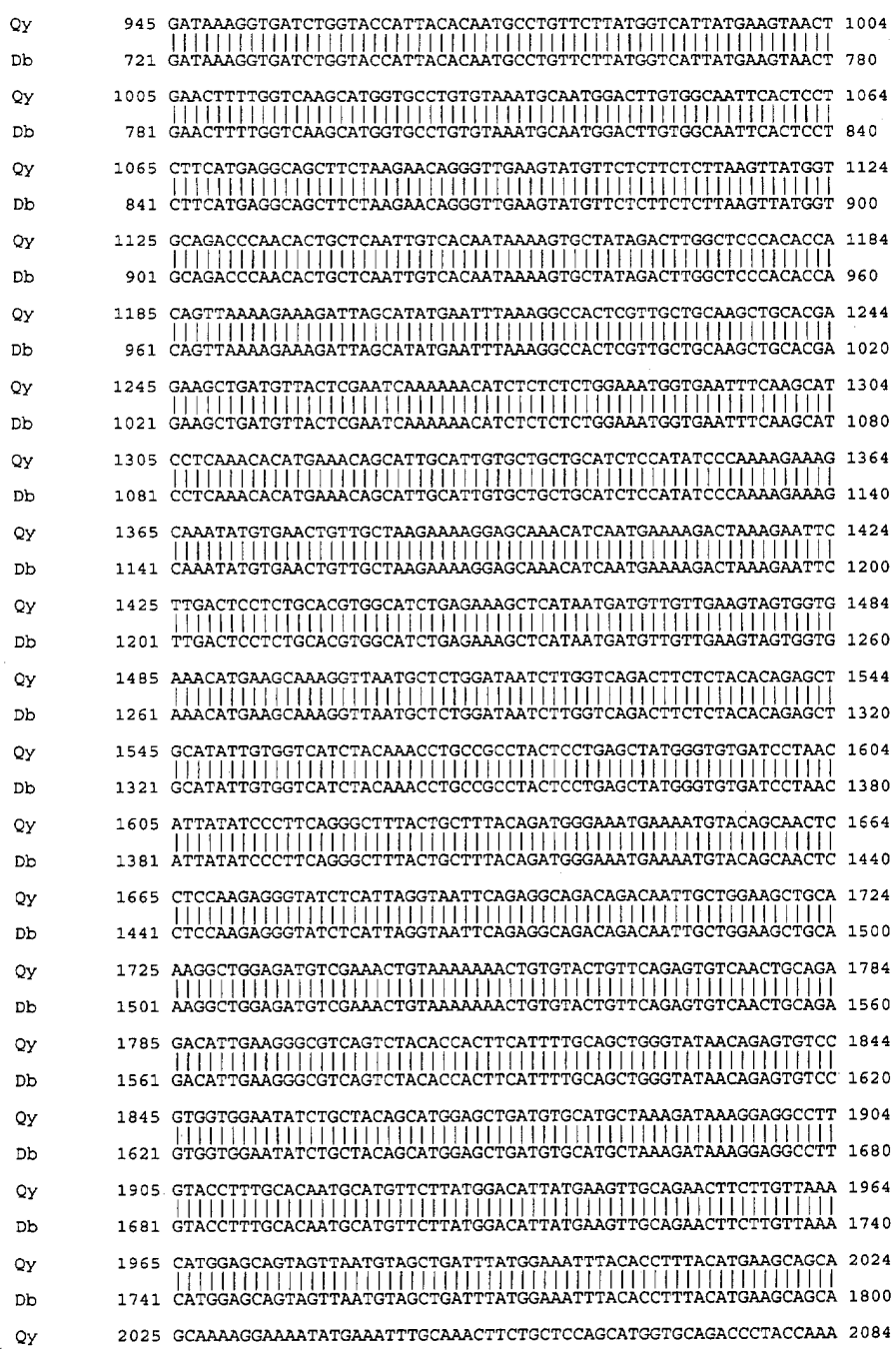
Qy        645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
Db        421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480

Qy        705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 764
Db        481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAA 540

Qy        765 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 824
Db        541 CTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTA 600

Qy        825 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 884
Db        601 AATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGA 660

Qy        885 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 944
Db        661 TATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAA 720
```


Qy 945 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 1004
Db 721 GATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACT 780
Qy 1005 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 1064
Db 781 GAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCT 840
Qy 1065 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 1124
Db 841 CTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGT 900
Qy 1125 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 1184
Db 901 GCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCA 960
Qy 1185 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1244
Db 961 CAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGA 1020
Qy 1245 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1304
Db 1021 GAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCAT 1080
Qy 1305 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1364
Db 1081 CCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAG 1140
Qy 1365 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1424
Db 1141 CAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTC 1200
Qy 1425 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1484
Db 1201 TTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTG 1260
Qy 1485 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1544
Db 1261 AAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCT 1320
Qy 1545 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1604
Db 1321 GCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAAC 1380
Qy 1605 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1664
Db 1381 ATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTC 1440
Qy 1665 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1724
Db 1441 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1500
Qy 1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
Db 1501 AAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1560
Qy 1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
Db 1561 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1620
Qy 1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
Db 1621 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1680
Qy 1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
Db 1681 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1740
Qy 1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
Db 1741 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 1800
Qy 2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2084
Db 1801 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAA 1860
Qy 2085 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2144
Db 1861 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 1920
Qy 2145 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2204
Db 1921 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 1980
Qy 2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
Db 1981 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2040
Qy 2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
Db 2041 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2100
Qy 2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
Db 2101 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2160
Qy 2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
Db 2161 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2220
Qy 2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
Db 2221 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2280
Qy 2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
Db 2281 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2340
Qy 2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
Db 2341 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2400
Qy 2625 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
Db 2401 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2460
Qy 2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
Db 2461 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2520
Qy 2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
Db 2521 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2580
Qy 2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
Db 2581 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2640
Qy 2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
Db 2641 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2700
Qy 2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
Db 2701 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2760
Qy 2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
Db 2761 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 2820
Qy 3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
Db 2821 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 2880
Qy 3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164

```
Db       2881 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 2940

Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3000

Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3060

Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3120

Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3180

Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3240

Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3300

Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3360

Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3420

Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3480

Qy       3705 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3764
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3540

Qy       3765 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3589
```

RESULT 9
US-10-199-937-1
; Sequence 1, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(3508)

US-10-199-937-1

Query Match 91.8%; Score 3504.8; DB 14; Length 3508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3506; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        215 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600

Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840

Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900

Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960

Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020
```

```
Qy      1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080

Qy      1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140

Qy      1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200

Qy      1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260

Qy      1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320

Qy      1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380

Qy      1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440

Qy      1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500

Qy      1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGT 1774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGT 1560

Qy      1775 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 CAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAA 1620

Qy      1835 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 CAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAA 1680

Qy      1895 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 AGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACT 1740

Qy      1955 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 2014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACA 1800

Qy      2015 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 2074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 TGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGA 1860

Qy      2075 CCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 2134
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 CCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATAC 1920

Qy      2135 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 2194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 AGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTG 1980

Qy      2195 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 TTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGG 2040

Qy      2255 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 CAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA 2100

Qy      2315 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 TTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACA 2160

Qy      2375 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2161 TAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATG 2220

Qy      2435 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2221 TGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACG 2280

Qy      2495 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2281 AACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGA 2340

Qy      2555 AGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGC 2614
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db      2341 AGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGC 2400

Qy      2615 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2401 CATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAG 2460

Qy      2675 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2734
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2461 CCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGC 2520

Qy      2735 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2521 AGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTC 2580

Qy      2795 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2581 AAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAG 2640

Qy      2855 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2914
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2641 CATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGA 2700

Qy      2915 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2701 ACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAAT 2760

Qy      2975 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 3034
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 CAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACA 2820

Qy      3035 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 3094
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2821 ACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGA 2880

Qy      3095 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 3154
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2881 TCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCG 2940

Qy      3155 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2941 AGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGAT 3000

Qy      3215 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3001 TCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3060

Qy      3275 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3061 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3120

Qy      3335 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3121 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3180

Qy      3395 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3454
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3240

Qy       3455 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3514
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3300

Qy       3515 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3574
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3360

Qy       3575 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3634
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3420

Qy       3635 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3694
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3480

Qy       3695 TATGAGGCCTGAAGGTATGGTCGATGGA 3722
                ||||||||||||||||||||||||||||
Db       3481 TATGAGGCCTGAAGGTATGGTCGATGGA 3508


RESULT 10
US-09-972-115A-3
; Sequence 3, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 4297
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3801)
;   OTHER INFORMATION:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (125)..(125)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (293)..(293)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (328)..(328)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (341)..(341)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (378)..(378)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (415)..(415)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (432)..(432)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (459)..(459)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (496)..(496)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (517)..(517)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (534)..(534)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (553)..(553)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (562)..(562)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (614)..(614)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (619)..(619)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (632)..(632)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (637)..(637)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (643)..(643)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (680)..(680)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (682)..(682)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (711)..(713)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (1971)..(1971)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2000)..(2000)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
```

```
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2175)..(2175)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2180)..(2180)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2205)..(2205)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2221)..(2221)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2289)..(2289)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2316)..(2316)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2359)..(2674)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2977)..(2977)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2291)..(2291)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
US-09-972-115A-3

  Query Match            87.6%;  Score 3341;  DB 10;  Length 4297;
  Best Local Similarity  89.0%;  Pred. No. 0;
  Matches 3400;  Conservative    8;  Mismatches  402;  Indels   12;  Gaps    2;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||  |  |||||||||||||||||||||||||||||||||||| |||  ||
Db         74 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCATCGAANGGGCGCGC 133

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
                ||    | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GTGGGCGCGGCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 193

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 253

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        254 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCANGATCATGTCGGGTCGCCGCT 313

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              |||||||||||||| |||||||||||| |||||||||||| |||||||| ||||||||||
Db        314 GCGCCGGCGGGGGANCGGCCTGCGCGANCGCCGCGGCCGAAGCCGTGGAACCGGCCGCCC 373

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              || | |||||||| |||||||||||||||||||||||||| |||||| |||||||||| |
Db        374 GAAANCTGTTCGAAGCGTGCCGCAACGGGGACGTGGAACGANTCAAGAAGCTGGTGACNC 433

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||:||||||||||||||||||| ||||||||||||||||||||||||||||: |||
Db        434 CTGARAAGGTGAACAGCCGCGACACNGCGGGCAGGAAATCCACCCCGCTGCACTTYCCCG 493

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              || |||||||||||||||| || ||:|||||||||||| ||||||||||||||:|||

Db        494 CANGTTTTGGGCGGAAAGACTTANTTRAATATTTGCTTCANAATGGTGCAAATGTYCAAN 553

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||:||||:||
Db        554 CACGTGATNATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCMTGCTRAAA 613

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600
                || |||||||||||| |||| ||||| |:|||||||||||||||||||||||||||||
Db        614 NTATCNATCTCCTTTTGCNACATNGTGCANAMCCCAATGCTCGAGATAATTGGAATTATA 673

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660
              |||||| | ||||||||||||||||||||||||||||| |||||||||||||||||||
Db        674 CTCCTCNCNATGAAGCTGCAATTAAAGGAAAGATTGANNNTTGCATTGTGCTGTTACAGC 733

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 793

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 853

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 913

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 973

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 1033

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1093

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1153

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1213

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1273

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1333

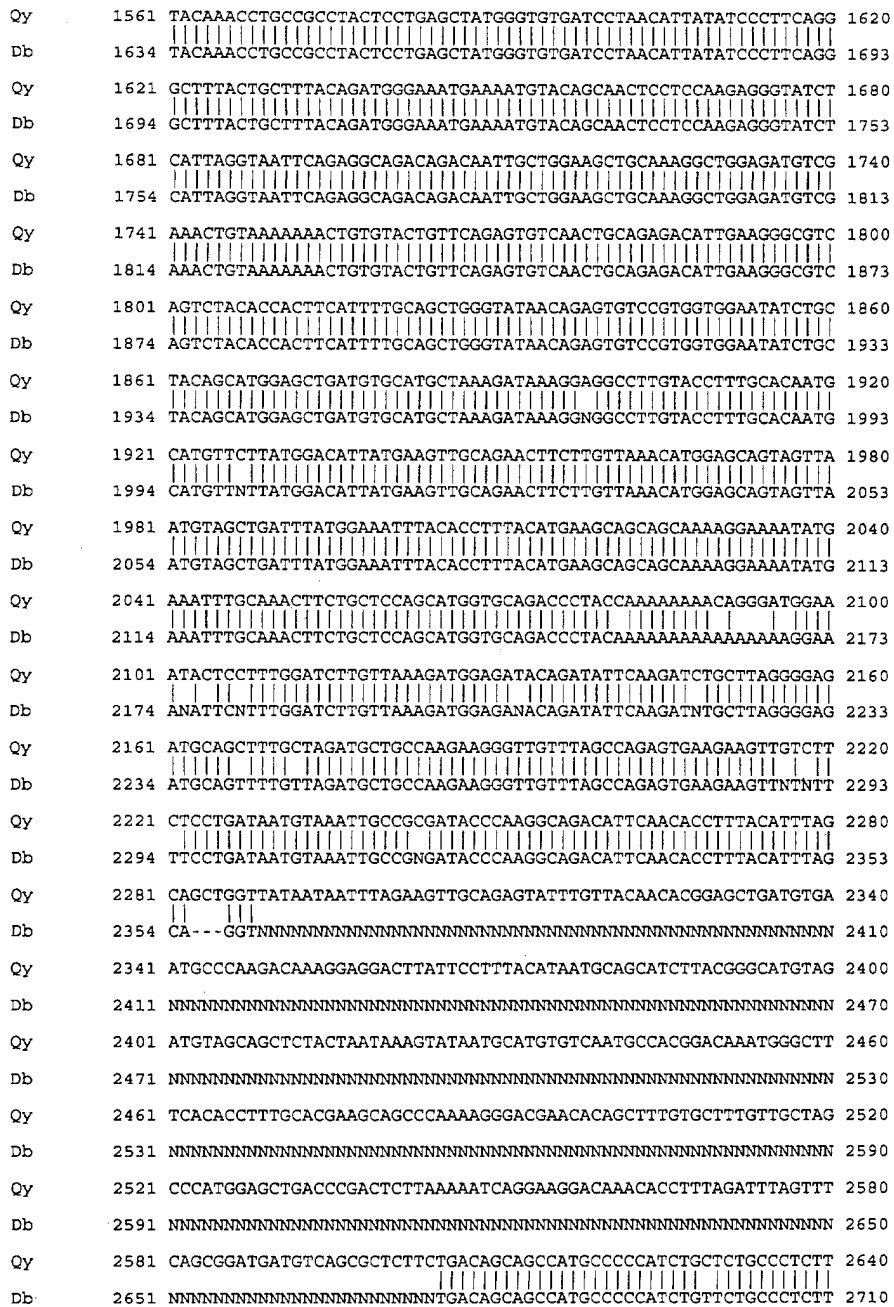
Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1393

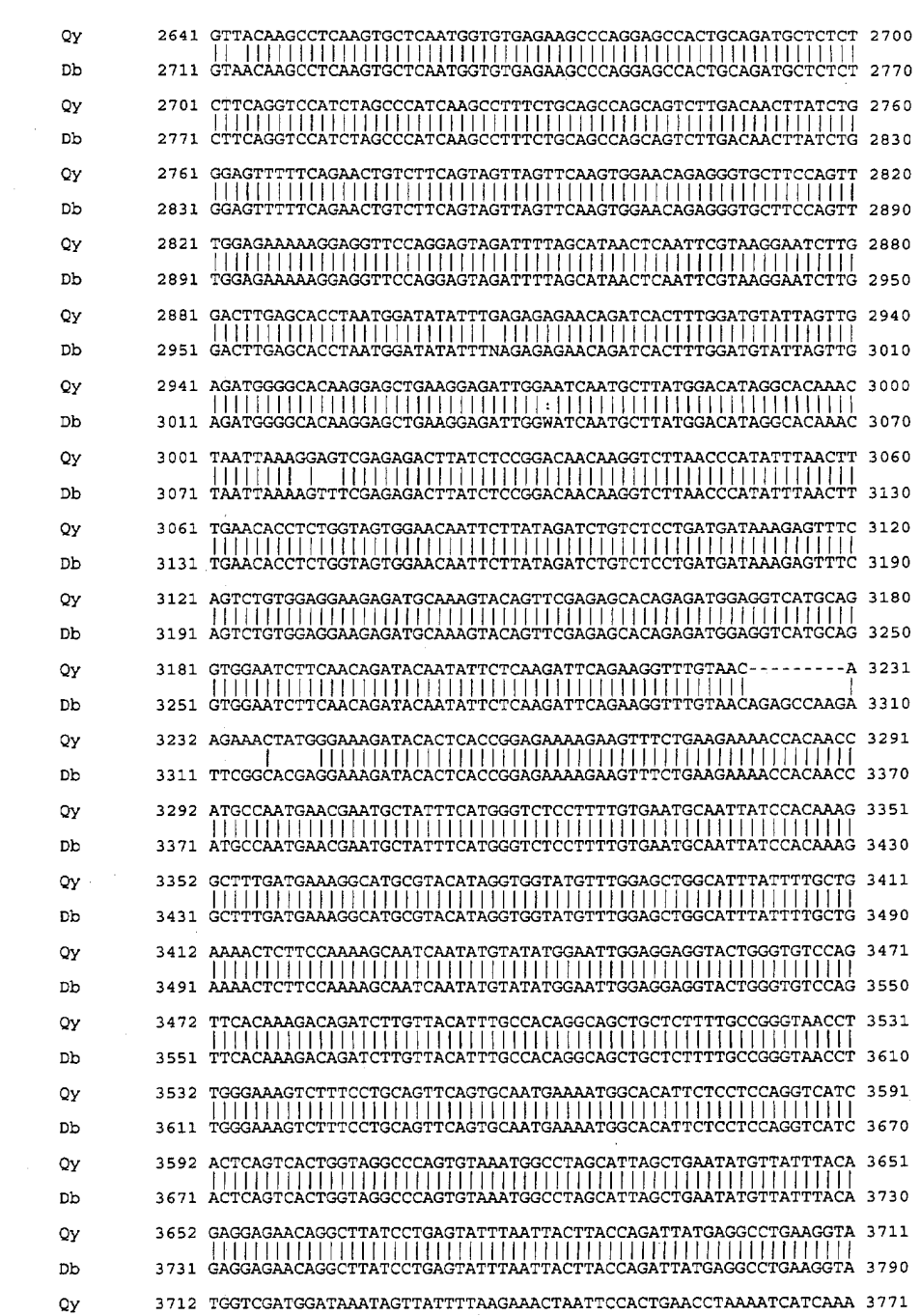
Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1453

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1513

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1573

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1633
```





```
Db       3791 TGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAA 3850

Qy       3772 GCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       3851 GCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3892


RESULT 11
US-09-509-196A-1
; Sequence 1, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
;  APPLICANT:  DALY, Roger J.
;  APPLICANT:  SUTHERLAND, Robert L.
;  TITLE OF INVENTION:  A Potential Effector for the Grb7 Family of Signalling
;  TITLE OF INVENTION:  Proteins
;  FILE REFERENCE:  1871-129
;  CURRENT APPLICATION NUMBER: US/09/509,196A
;  CURRENT FILING DATE:  2000-03-23
;  PRIOR APPLICATION NUMBER:  P09388
;  PRIOR FILING DATE:  1997-09-23
;  PRIOR APPLICATION NUMBER:  PCT AU98/00795
;  PRIOR FILING DATE:  1998-09-23
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 3400
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-509-196A-1

  Query Match             86.4%;  Score 3297;  DB 9;  Length 3400;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        501 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 560
Db          1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60

Qy        561 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 620
Db         61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120

Qy        621 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 680
Db        121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180

Qy        681 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 740
Db        181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240

Qy        741 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 800
Db        241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300

Qy        801 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 860
Db        301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360

Qy        861 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 920
Db        361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420

Qy        921 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 980
Db        421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480

Qy        981 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1040
Db        481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540

Qy       1041 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1100
Db        541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600

Qy       1101 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1160
Db        601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660

Qy       1161 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1220
Db        661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720

Qy       1221 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1280
Db        721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780

Qy       1281 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1340
Db        781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840

Qy       1341 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1400
Db        841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900

Qy       1401 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1460
Db        901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960

Qy       1461 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1520
Db        961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020

Qy       1521 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1580
Db       1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080

Qy       1581 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1640
Db       1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy       1641 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1700
Db       1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200

Qy       1701 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1760
Db       1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260

Qy       1761 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1820
Db       1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320

Qy       1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1880
Db       1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380

Qy       1881 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1940
Db       1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440

Qy       1941 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 2000
Db       1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500

Qy       2001 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2060
Db       1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560

Qy       2061 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2120
Db       1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620

Qy       2121 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2180
```

```
Db       1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680
Qy       2181 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2240
Db       1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740
Qy       2241 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2300
Db       1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800
Qy       2301 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2360
Db       1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860
Qy       2361 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2420
Db       1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920
Qy       2421 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2480
Db       1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980
Qy       2481 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2540
Db       1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040
Qy       2541 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2600
Db       2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100
Qy       2601 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2660
Db       2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160
Qy       2661 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2720
Db       2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220
Qy       2721 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2780
Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280
Qy       2781 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2840
Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340
Qy       2841 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2900
Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400
Qy       2901 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2960
Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460
Qy       2961 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3020
Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520
Qy       3021 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3080
Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580
Qy       3081 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3140
Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640
Qy       3141 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3200
Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700
Qy       3201 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3260
Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3261 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3320
Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820
Qy       3321 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3380
Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880
Qy       3381 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3440
Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940
Qy       3441 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3500
Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000
Qy       3501 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3560
Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060
Qy       3561 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3620
Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120
Qy       3621 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3680
Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180
Qy       3681 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3740
Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240
Qy       3741 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3800
Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300
Qy       3801 TGCTGAAAAAAAA 3813
Db       3301 TGCTGAAAAAAAA 3313
```

RESULT 12
US-09-972-115A-1
; Sequence 1, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3999)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (1)..(586)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.

; NAME/KEY: misc_feature
; LOCATION: (596)..(597)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (621)..(621)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (626)..(626)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (629)..(629)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (638)..(638)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (693)..(694)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (724)..(724)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (730)..(730)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (764)..(764)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (791)..(791)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2148)..(2148)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2177)..(2177)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2382)..(2382)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2466)..(2466)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2468)..(2468)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2493)..(2493)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2536)..(2860)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature

; LOCATION: (3163)..(3163)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (3488)..(3499)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.

US-09-972-115A-1

Query Match 79.6%; Score 3039; DB 10; Length 4493;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 3101; Conservative 1; Mismatches 375; Indels 27; Gaps 2;

```
Qy      337 AACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGA 396
            ||| | ||   |||||||||| ||||| ||||| |||| || | |||||| ||||||||
Db      587 AACCATTCCNNAGGCTGGTGACCCCTGAAAAGGTNAACANCCNCAACACGGNGGGCAGGA 646

Qy      397 AATCCACCCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGC 456
            |||||||||||||||||| |||||||||||||||||| || | || |||||||||
Db      647 AATCCACCCCGCTGCACTTCCCCGCAGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGC 706

Qy      457 TTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATG 516
            |||| |||||||||||| ||||| ||| | || ||||||||||||||||||||||| ||
Db      707 TTCAAAATGGTGCAAATNTCCAANCACTTTATAATGGGGGCCTTATTCCTCTTCATANTG 766

Qy      517 CATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCA 576
            |||||||||||||||||| ||     ||||||||||||||||||||||||||||||||
Db      767 CATGCTCTTTTGGTCATGCTAAAANTATCAATCTCCTTTTGCGACATGGTGCAGACCCCA 826

Qy      577 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 636
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      827 ATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTG 886

Qy      637 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 696
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      887 ATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAA 946

Qy      697 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 756
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      947 GGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGA 1006

Qy      757 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 816
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1007 AAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCA 1066

Qy      817 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 876
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1067 CACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGG 1126

Qy      877 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 936
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1127 CAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCC 1186

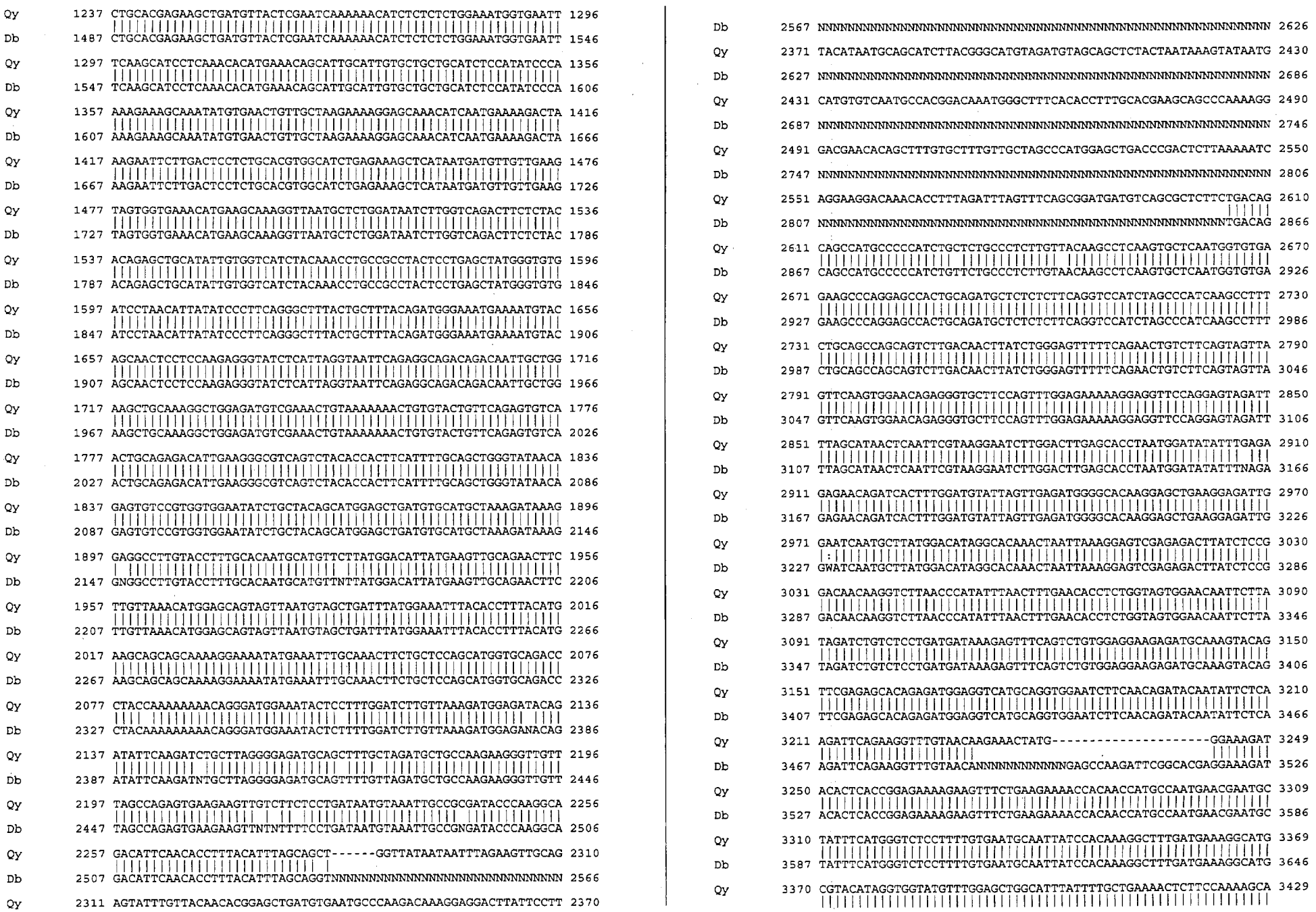
Qy      937 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 996
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1187 ATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATG 1246

Qy      997 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1056
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1247 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1306

Qy     1057 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1116
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1307 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1366

Qy     1117 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1176
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1367 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1426

Qy     1177 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1236
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1427 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1486
```


Qy 1237 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1296
Db 1487 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1546
Qy 1297 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1356
Db 1547 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1606
Qy 1357 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1416
Db 1607 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1666
Qy 1417 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1476
Db 1667 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1726
Qy 1477 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1536
Db 1727 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1786
Qy 1537 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1596
Db 1787 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1846
Qy 1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656
Db 1847 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1906
Qy 1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716
Db 1907 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1966
Qy 1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776
Db 1967 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 2026
Qy 1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836
Db 2027 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 2086
Qy 1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896
Db 2087 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 2146
Qy 1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956
Db 2147 GNGGCCTTGTACCTTTGCACAATGCATGTTNTTATGGACATTATGAAGTTGCAGAACTTC 2206
Qy 1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2016
Db 2207 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2266
Qy 2017 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2076
Db 2267 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2326
Qy 2077 CTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2136
Db 2327 CTACAAAAAAAAACAGGGATGGAAATACTCTTTTGGATCTTGTTAAAGATGGAGANACAG 2386
Qy 2137 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2196
Db 2387 ATATTCAAGATNTGCTTAGGGGAGATGCAGTTTTGTTAGATGCTGCCAAGAAGGGTTGTT 2446
Qy 2197 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2256
Db 2447 TAGCCAGAGTGAAGAAGTTNTNTTTTCCTGATAATGTAAATTGCCGNGATACCCAAGGCA 2506
Qy 2257 GACATTCAACACCTTTACATTTAGCAGCT------GGTTATAATAATTTAGAAGTTGCAG 2310
Db 2507 GACATTCAACACCTTTACATTTAGCAGGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2566
Qy 2311 AGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTT 2370
Db 2567 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2626
Qy 2371 TACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATG 2430
Db 2627 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2686
Qy 2431 CATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGG 2490
Db 2687 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2746
Qy 2491 GACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATC 2550
Db 2747 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2806
Qy 2551 AGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAG 2610
Db 2807 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAG 2866
Qy 2611 CAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGA 2670
Db 2867 CAGCCATGCCCCCATCTGTTCTGCCCTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGA 2926
Qy 2671 GAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTT 2730
Db 2927 GAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTT 2986
Qy 2731 CTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTA 2790
Db 2987 CTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTA 3046
Qy 2791 GTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATT 2850
Db 3047 GTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATT 3106
Qy 2851 TTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGA 2910
Db 3107 TTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTNAGA 3166
Qy 2911 GAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTG 2970
Db 3167 GAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTG 3226
Qy 2971 GAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCG 3030
Db 3227 GWATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCG 3286
Qy 3031 GACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTA 3090
Db 3287 GACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTA 3346
Qy 3091 TAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAG 3150
Db 3347 TAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAG 3406
Qy 3151 TTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCA 3210
Db 3407 TTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCA 3466
Qy 3211 AGATTCAGAAGGTTTGTAACAAGAAACTATG----------GGAAAGAT 3249
Db 3467 AGATTCAGAAGGTTTGTAACANNNNNNNNNNNGAGCCAAGATTCGGCACGAGGAAAGAT 3526
Qy 3250 ACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGC 3309
Db 3527 ACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGC 3586
Qy 3310 TATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATG 3369
Db 3587 TATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATG 3646
Qy 3370 CGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCA 3429

```
Db       3647 CGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCA 3706

Qy       3430 ATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTT 3489
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3707 ATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTT 3766

Qy       3490 GTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGC 3549
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3767 GTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGC 3826

Qy       3550 AGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGC 3609
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3827 AGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGC 3886

Qy       3610 CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATC 3669
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3887 CCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATC 3946

Qy       3670 CTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAG 3729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3947 CTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAG 4006

Qy       3730 TTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACG 3789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4007 TTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACG 4066

Qy       3790 TTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||
Db       4067 TTTTACTCCTTTGCTGAAAAAAAA 4090
```

RESULT 13
US-10-199-937-92
; Sequence 92, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 2971
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-937-92

Query Match 63.9%; Score 2439.8; DB 14; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1371 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1430
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1431 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1490
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1491 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1551 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1670
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1671 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1730
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1731 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1790
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1791 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1850
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1851 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1910
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1911 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1970
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1971 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2030
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2031 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2090
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAAC 720

Qy       2091 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2151 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2210
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2211 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2270
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2271 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2330
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2331 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2390
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2450
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2451 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2510
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy       2511 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2570
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy       2571 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630
             ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2631 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2690
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320
Qy       2691 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2750
Db       1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380
Qy       2751 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2810
Db       1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440
Qy       2811 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2870
Db       1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500
Qy       2871 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930
Db       1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560
Qy       2931 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2990
Db       1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620
Qy       2991 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3050
Db       1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680
Qy       3051 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3110
Db       1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740
Qy       3111 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3170
Db       1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800
Qy       3171 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3230
Db       1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860
Qy       3231 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3290
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920
Qy       3291 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3350
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980
Qy       3351 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3410
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040
Qy       3411 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3470
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100
Qy       3471 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3530
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160
Qy       3531 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3590
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220
Qy       3591 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3650
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280
Qy       3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340
Qy       3711 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3770
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400
Qy       3771 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 14
US-10-199-937-88
; Sequence 88, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 3353
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2352)
US-10-199-937-88

Query Match 63.9%; Score 2439.8; DB 14; Length 3353;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1371 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1430
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60
Qy       1431 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1490
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120
Qy       1491 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1550
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180
Qy       1551 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1610
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240
Qy       1611 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1670
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300
Qy       1671 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1730
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360
Qy       1731 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1790
Db        361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420
Qy       1791 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1850
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480
Qy       1851 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1910
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540
```

```
Qy       1911 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1970
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1971 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2030
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2031 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAAC 2090
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAAC 720

Qy       2091 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2150
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2151 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2210
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2211 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2270
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2271 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2330
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2331 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2390
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2391 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2450
Db        1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2451 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2510
Db        1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy       2511 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2570
Db        1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy       2571 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2630
Db        1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2631 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2690
Db        1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy       2691 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2750
Db        1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy       2751 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2810
Db        1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy       2811 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2870
Db        1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy       2871 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2930
Db        1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy       2931 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2990
Db        1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2991 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3050
Db        1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy       3051 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3110
Db        1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy       3111 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3170
Db        1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy       3171 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3230
Db        1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860

Qy       3231 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3290
Db        1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy       3291 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3350
Db        1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy       3351 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3410
Db        1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy       3411 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3470
Db        2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy       3471 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3530
Db        2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy       3531 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3590
Db        2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3591 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3650
Db        2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3651 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3710
Db        2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3711 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3770
Db        2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy       3771 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
Db        2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 15
US-10-199-937-90
; Sequence 90, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582

```
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
;   LENGTH: 3799
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(2270)
US-10-199-937-90

  Query Match             61.8%;  Score 2357.8;  DB 14;  Length 3799;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1453 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 1512
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 60

Qy       1513 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 1572
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 120

Qy       1573 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 1632
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 180

Qy       1633 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 1692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 240

Qy       1693 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1752
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 300

Qy       1753 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 1812
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 360

Qy       1813 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 1872
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 420

Qy       1873 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 1932
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 480

Qy       1933 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 1992
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 540

Qy       1993 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 2052
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 600

Qy       2053 TTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGG 2112
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        601 TTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGG 660

Qy       2113 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 2172
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 720

Qy       2173 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 2232
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 780

Qy       2233 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 2292
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 840

Qy       2293 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 900

Qy       2353 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 960

Qy       2413 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 1020

Qy       2473 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 1080

Qy       2533 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2592
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        1081 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 1140

Qy       2593 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2652
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 1200

Qy       2653 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2712
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 1260

Qy       2713 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2772
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 1320

Qy       2773 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2832
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 1380

Qy       2833 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2892
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 1440

Qy       2893 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2952
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 1500

Qy       2953 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 3012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 1560

Qy       3013 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 3072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 1620

Qy       3073 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 3132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 1680

Qy       3133 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 3192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 1740

Qy       3193 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 3252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 1800

Qy       3253 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 3312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 1860

Qy       3313 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 3372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 1920

Qy       3373 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 3432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

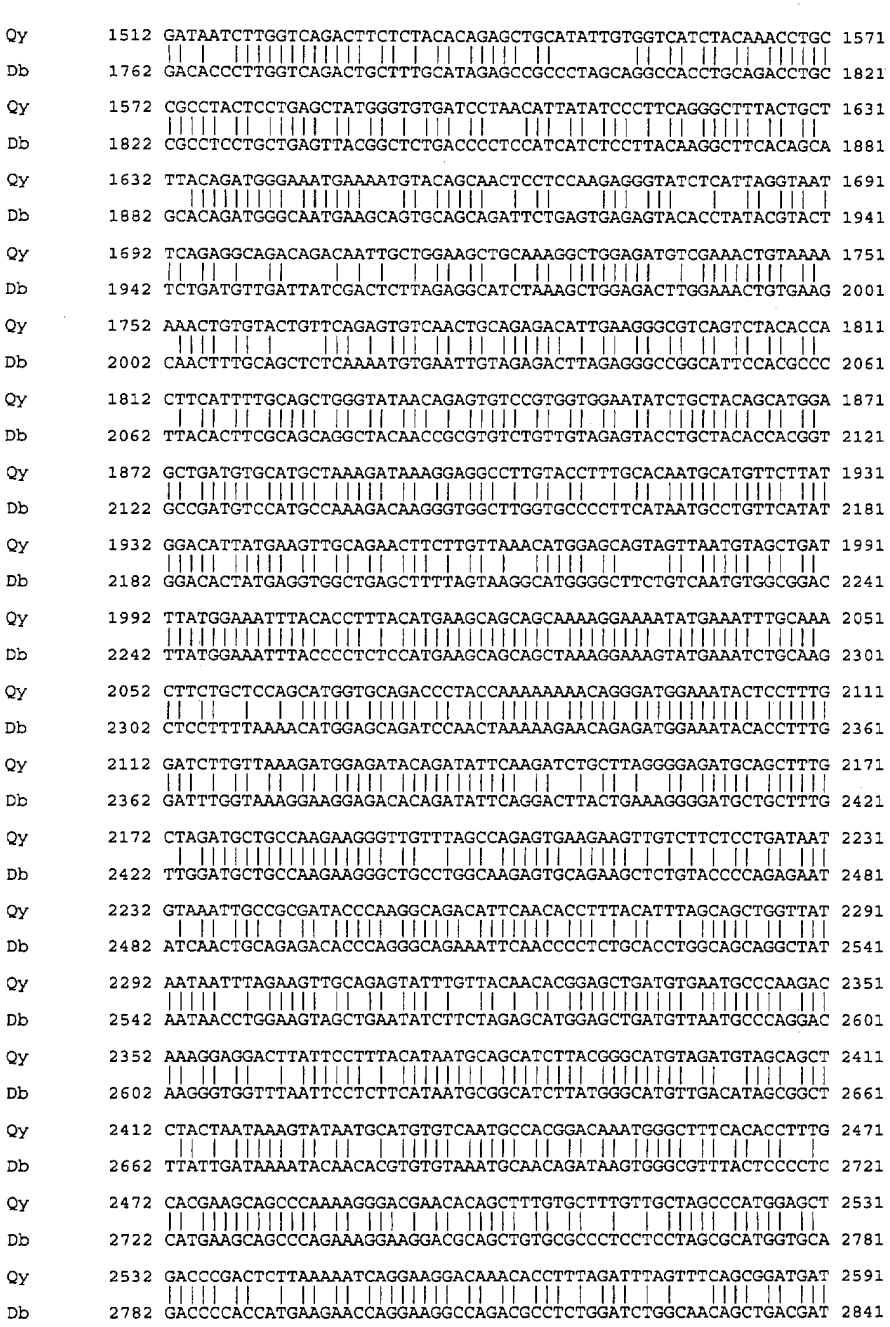
```
Db       1921 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 1980
Qy       3433 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 3492
Db       1981 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 2040
Qy       3493 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 3552
Db       2041 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 2100
Qy       3553 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 3612
Db       2101 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 2160
Qy       3613 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 3672
Db       2161 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 2220
Qy       3673 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 3732
Db       2221 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 2280
Qy       3733 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 3792
Db       2281 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 2340
Qy       3793 TACTCCTTTGCTGAAAAAAAA 3813
Db       2341 TACTCCTTTGCTGAAAAAAAA 2361
```

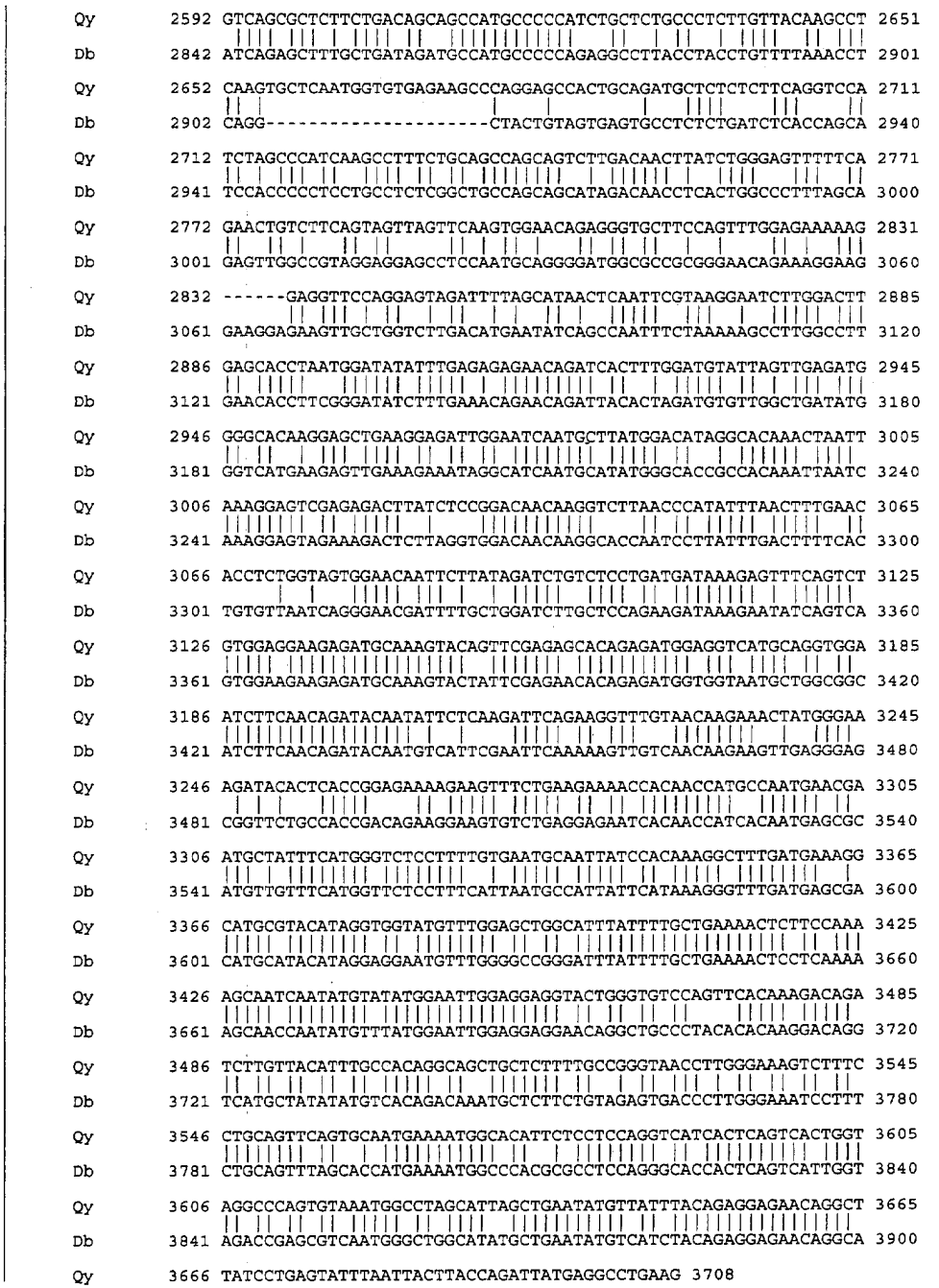
RESULT 16
US-10-199-937-3
; Sequence 3, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3981)
US-10-199-937-3

Query Match 49.4%; Score 1885; DB 14; Length 3984;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps 2;

```
Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db        502 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 561
Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        562 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 621
Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
```

```
Db        622 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 681
Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        682 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 741
Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        742 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 801
Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        802 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 861
Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        862 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 921
Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        922 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 981
Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        982 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1041
Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1042 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1101
Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1102 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1161
Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1162 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1221
Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db       1222 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1281
Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db       1282 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1341
Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db       1342 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1401
Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db       1402 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1461
Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db       1462 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1521
Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db       1522 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1581
Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
Db       1582 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1641
Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       1642 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1701
Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       1702 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1761
```





```
Db       3901 TACCCAGAGTATCTTATCACTTACCAGATCATGAAGCCAGAAG 3943

RESULT 17
US-09-841-835-1
; Sequence 1, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4134 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-841-835-1

  Query Match           49.4%;  Score 1885;  DB 9;  Length 4134;
  Best Local Similarity   72.1%;  Pred. No. 0;
  Matches 2496;  Conservative    0;  Mismatches  940;  Indels   27;  Gaps    2;

Qy        252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db        507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy        312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db        567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy        372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db        627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy        432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db        687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy        492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db        747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy        552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db        807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy        612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db        867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy        672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db        927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy        732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db        987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy        792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db       1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy        852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db       1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy        912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db       1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy        972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db       1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy       1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db       1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy       1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db       1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy       1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db       1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy       1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db       1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy       1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db       1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

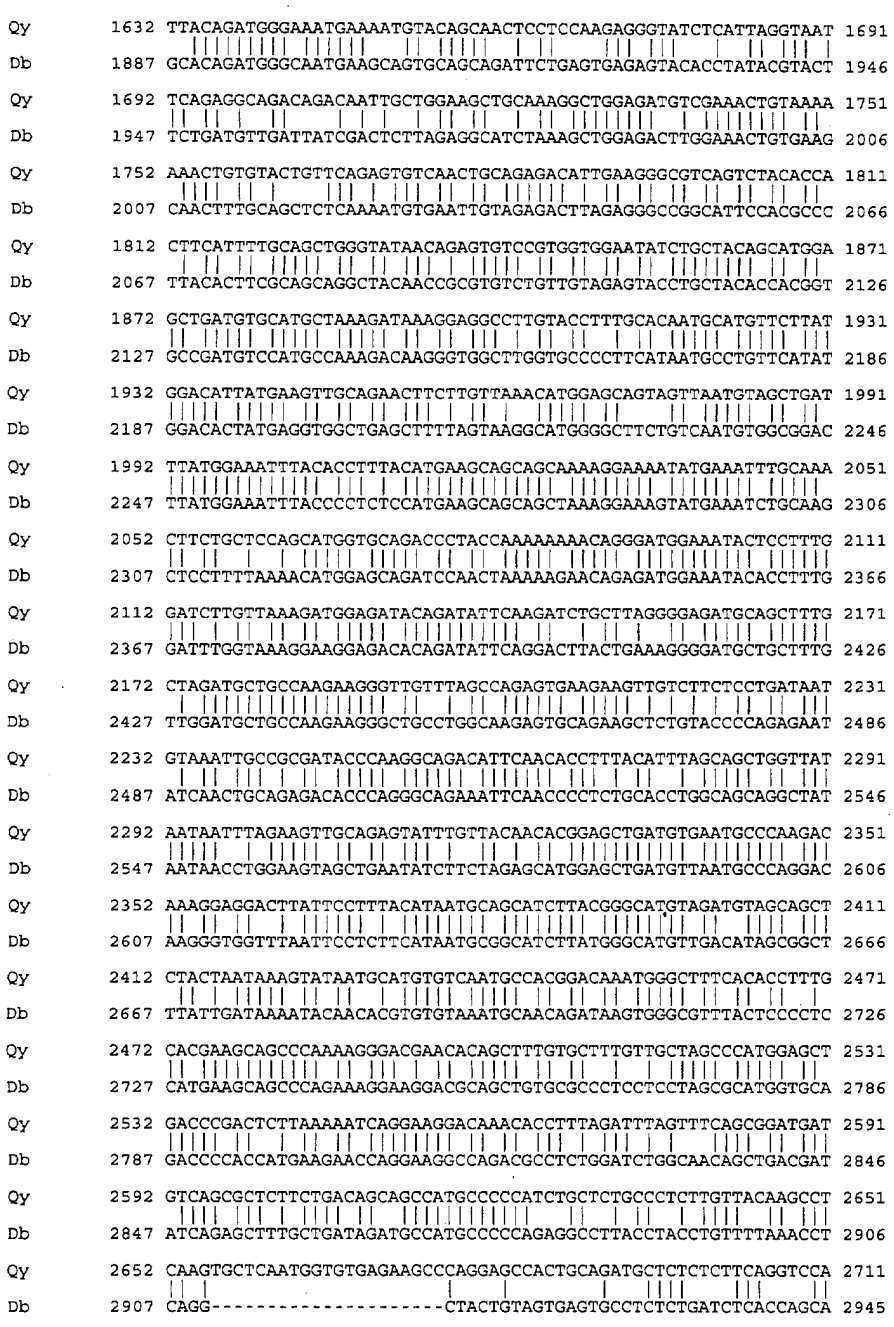
Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
Db       1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

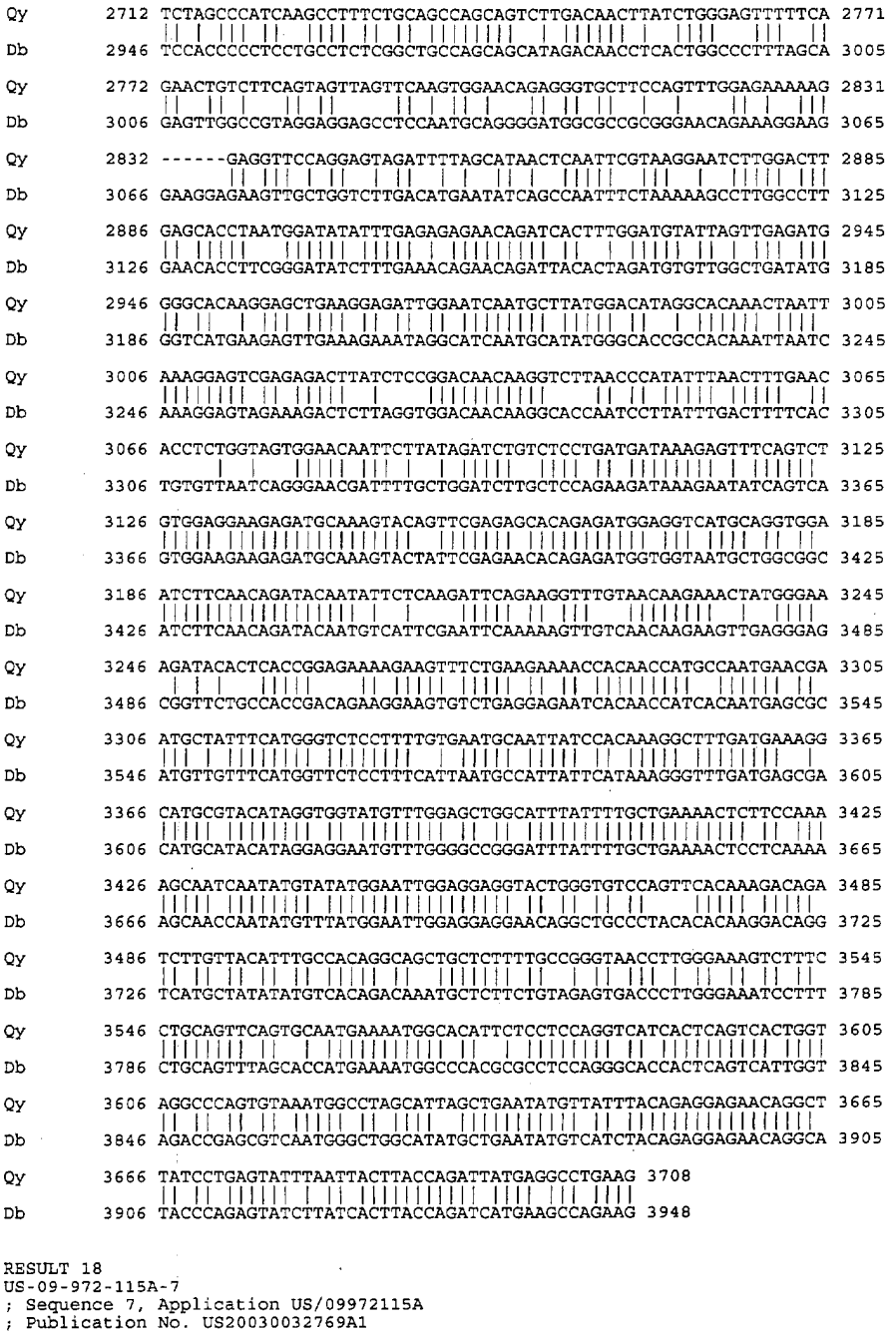
Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db       1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db       1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db       1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db       1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886
```

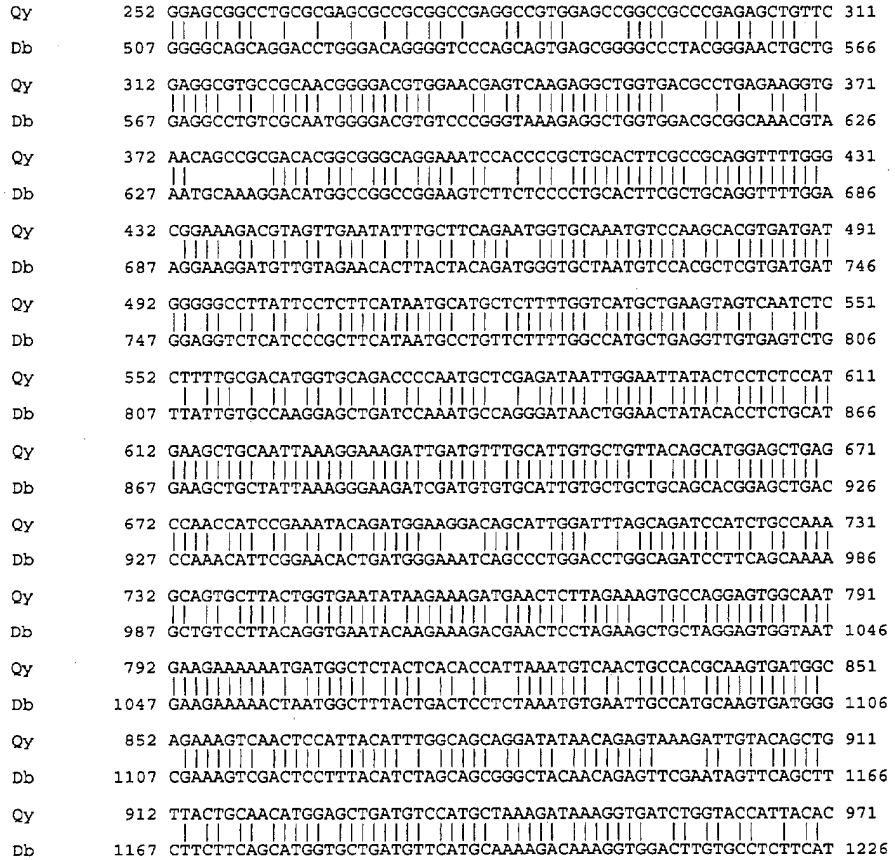


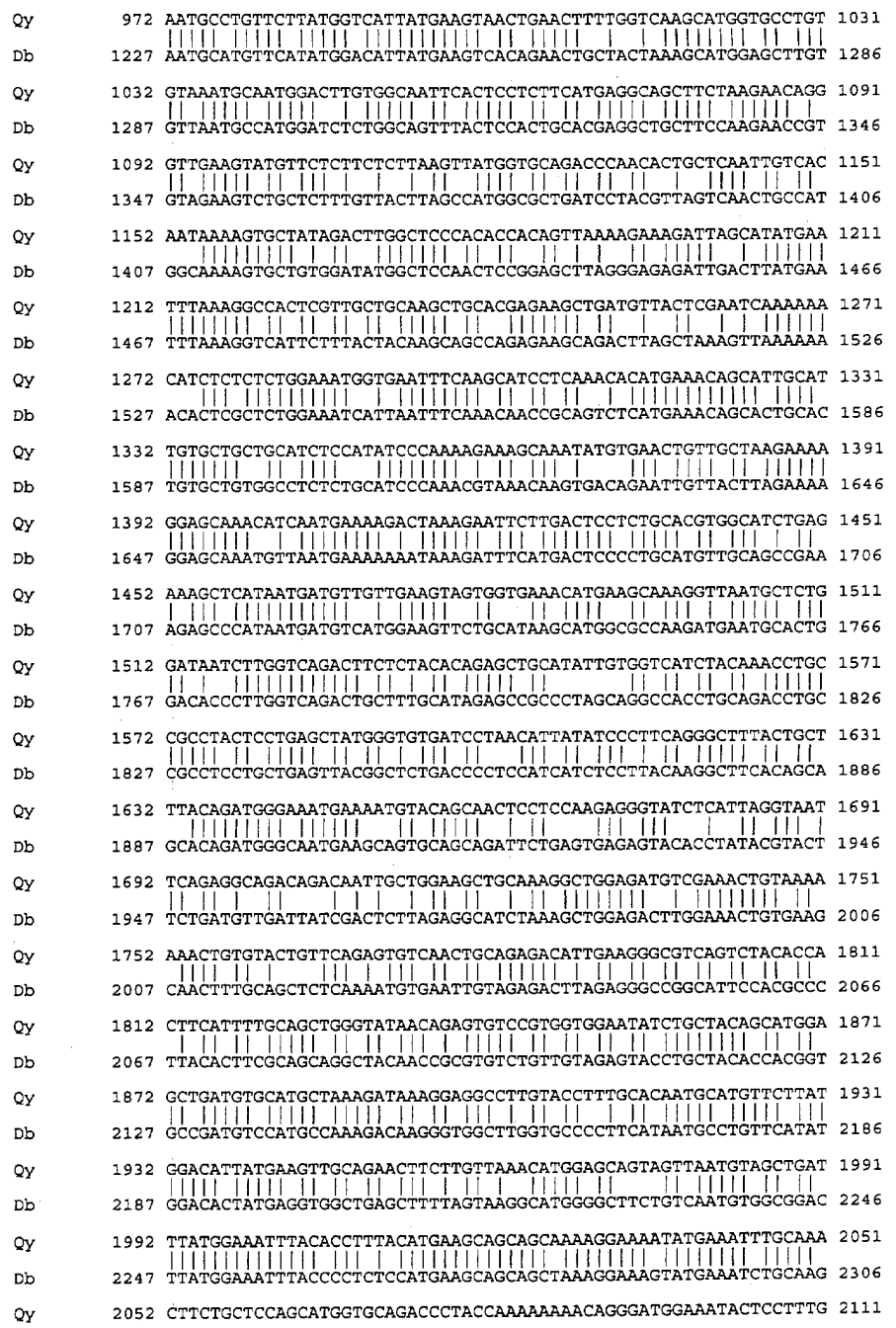


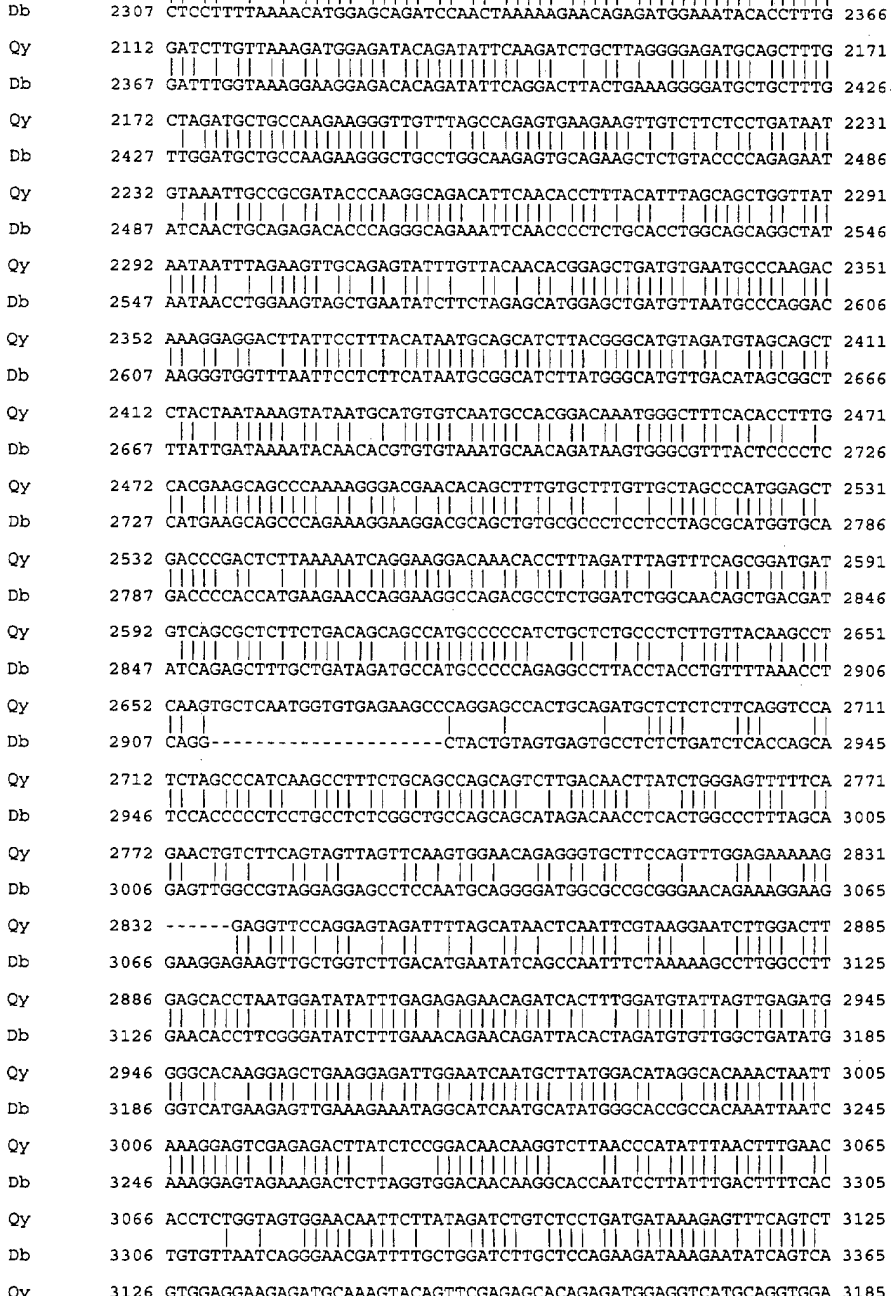
RESULT 18
US-09-972-115A-7
; Sequence 7, Application US/09972115A
; Publication No. US20030032769A1

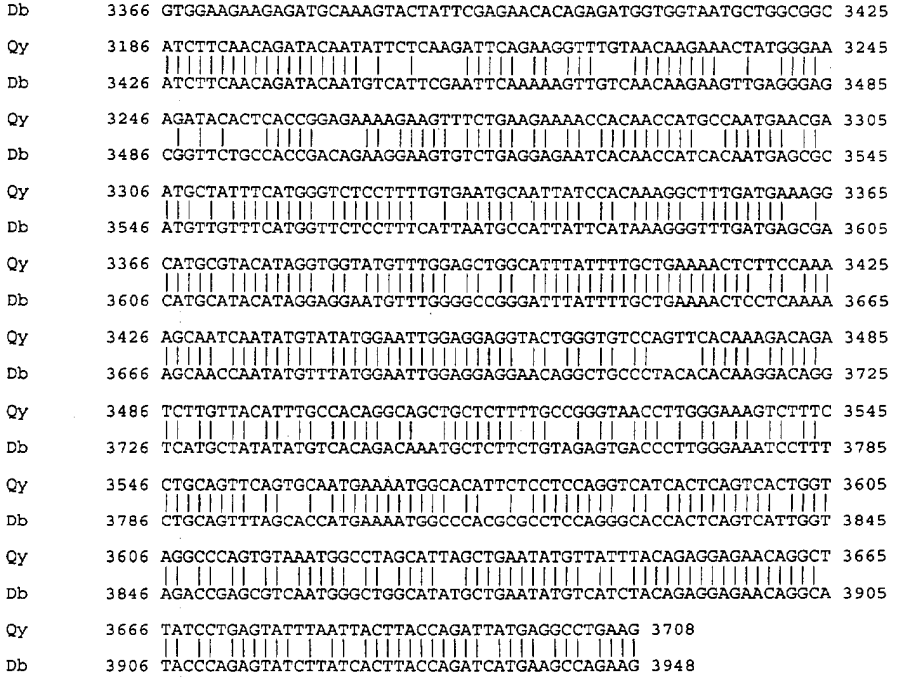
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 4134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-115A-7

Query Match 49.4%; Score 1885; DB 10; Length 4134;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 940; Indels 27; Gaps 2;









RESULT 19
US-10-199-937-98
; Sequence 98, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1567)
US-10-199-937-98

Query Match 41.1%; Score 1568; DB 14; Length 1568;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
Db          1 GGCCAGGATCATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGC 60

Qy        275 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGT 120

Qy        335 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAG 180

Qy        395 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTT 240

Qy        455 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAA 300

Qy        515 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 574
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCC 360

Qy        575 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGAT 420

Qy        635 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGG 480

Qy        695 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAA 540

Qy        755 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 814
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACT 600

Qy        815 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 874
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTT 660

Qy        875 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGT 720

Qy        935 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTA 780

Qy        995 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 1054
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCA 840

Qy       1055 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTT 900

Qy       1115 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGC 960

Qy       1175 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCA 1020

Qy       1235 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAA 1080

Qy       1295 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCC 1140

Qy       1355 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGAC 1200

Qy       1415 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1474
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGA 1260

Qy       1475 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCT 1320

Qy       1535 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1594
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTG 1380

Qy       1595 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGT 1440

Qy       1655 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1714
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCT 1500

Qy       1715 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGT 1560

Qy       1775 CAACTGCA 1782
              ||||||||
Db       1561 CAACTGCA 1568
```

RESULT 20
US-10-252-157-342
; Sequence 342, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 342
; LENGTH: 4435
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 474266.2
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4369, 4374
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-342

Query Match 40.2%; Score 1533.8; DB 14; Length 4435;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1562; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

```
Qy       2205 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2264
              || || |||| |  | | ||| | | |          | |  ||||  |   | | |
Db        289 GTTAACAAGTGGAAGTATTATGACATAGGACTAAAAAAGAAAAAACAAGTAAAGTAATTT 348

Qy       2265 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2324
                  ||   | |  ||||||||||||||||||||||||||||||||||||||||||||
Db        349 TACTTCTTTTCTAAATTAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 408
```

```
Qy       2325 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 468

Qy       2385 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2444
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 528

Qy       2445 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2504
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 588

Qy       2505 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2564
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 648

Qy       2565 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 708

Qy       2625 TCTGCTCTGCCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 768

Qy       2685 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2744
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 828

Qy       2745 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2804
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 888

Qy       2805 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 948

Qy       2865 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2924
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 1008

Qy       2925 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2984
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 1068

Qy       2985 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 1128

Qy       3045 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3104
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 1188

Qy       3105 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 1248

Qy       3165 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1249 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 1308

Qy       3225 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 1368

Qy       3285 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3344
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1369 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 1428

Qy       3345 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3404
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1429 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 1488

Qy       3405 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3464
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1489 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 1548

Qy       3465 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3524
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1549 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 1608

Qy       3525 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3584
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1609 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 1668

Qy       3585 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3644
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 1728

Qy       3645 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3704
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1729 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 1788

Qy       3705 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3764
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 1848

Qy       3765 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3813
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1849 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 1897
```

RESULT 21
US-09-841-835-7
; Sequence 7, Application US/09841835
; Patent No. US20020076795A1
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4491 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2027
US-09-841-835-7

  Query Match            39.9%;  Score 1521.2;  DB 9;  Length 4491;
  Best Local Similarity  65.4%;  Pred. No. 0;
  Matches 2498;  Conservative    0;  Mismatches  938;  Indels  384;  Gaps    3;

Qy       252 GGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTC 311
Db       507 GGGGCAGCAGGACCTGGGACAGGGGTCCCAGCAGTGAGCGGGGCCCTACGGGAACTGCTG 566

Qy       312 GAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTG 371
Db       567 GAGGCCTGTCGCAATGGGGACGTGTCCCGGGTAAAGAGGCTGGTGGACGCGGCAAACGTA 626

Qy       372 AACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGG 431
Db       627 AATGCAAAGGACATGGCCGGCCGGAAGTCTTCTCCCCTGCACTTCGCTGCAGGTTTTGGA 686

Qy       432 CGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGTGATGAT 491
Db       687 AGGAAGGATGTTGTAGAACACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGAT 746

Qy       492 GGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTC 551
Db       747 GGAGGTCTCATCCCGCTTCATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTG 806

Qy       552 CTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCAT 611
Db       807 TTATTGTGCCAAGGAGCTGATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCAT 866

Qy       612 GAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAG 671
Db       867 GAAGCTGCTATTAAAGGGAAGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGAC 926

Qy       672 CCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAA 731
Db       927 CCAAACATTCGGAACACTGATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAA 986

Qy       732 GCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAAT 791
Db       987 GCTGTCCTTACAGGTGAATACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAAT 1046

Qy       792 GAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGC 851
Db      1047 GAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGG 1106

Qy       852 AGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTG 911
Db      1107 CGAAAGTCGACTCCTTTACATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTT 1166

Qy       912 TTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACAC 971
Db      1167 CTTCTTCAGCATGGTGCTGATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCAT 1226

Qy       972 AATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGT 1031
Db      1227 AATGCATGTTCATATGGACATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGT 1286

Qy      1032 GTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGG 1091
Db      1287 GTTAATGCCATGGATCTCTGGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGT 1346

Qy      1092 GTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCAC 1151
Db      1347 GTAGAAGTCTGCTCTTTGTTACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCAT 1406

Qy      1152 AATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAA 1211
Db      1407 GGCAAAAGTGCTGTGGATATGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAA 1466

Qy      1212 TTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAA 1271
Db      1467 TTTAAAGGTCATTCTTTACTACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAA 1526

Qy      1272 CATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCAT 1331
Db      1527 ACACTCGCTCTGGAAATCATTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCAC 1586

Qy      1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
Db      1587 TGTGCTGTGGCCTCTCTGCATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAA 1646

Qy      1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
Db      1647 GGAGCAAATGTTAATGAAAAAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAA 1706

Qy      1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
Db      1707 AGAGCCCATAATGATGTCATGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTG 1766

Qy      1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
Db      1767 GACACCCTTGGTCAGACTGCTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGC 1826

Qy      1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
Db      1827 CGCCTCCTGCTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1886

Qy      1632 TTACAGATGGGAAATGAAAATGTACAGCAACT---------------------------- 1663
Db      1887 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGTGAGTTACGGCTCTGACCCC 1946

Qy      1664 ------------------------------------------------------------ 1663
Db      1947 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 2006

Qy      1664 ------------------------------------------------------------ 1663
Db      2007 ATTCTGAGTGGTCATTCGTAGATAGTGATCATTCTACTTCAGCCTTAATGGTGATCTTGA 2066

Qy      1664 ------------------------------------------------------------ 1663
Db      2067 GACGGGAAGATTTAGAAGGAAATCTATCCAGCATGTCTTCACTGTCAACATGAAGAGTAC 2126

Qy      1664 ------------------------------------------------------------ 1663
Db      2127 ACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTT 2186

Qy      1664 ------------------------------------------------------------ 1663
Db      2187 GGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCG 2246

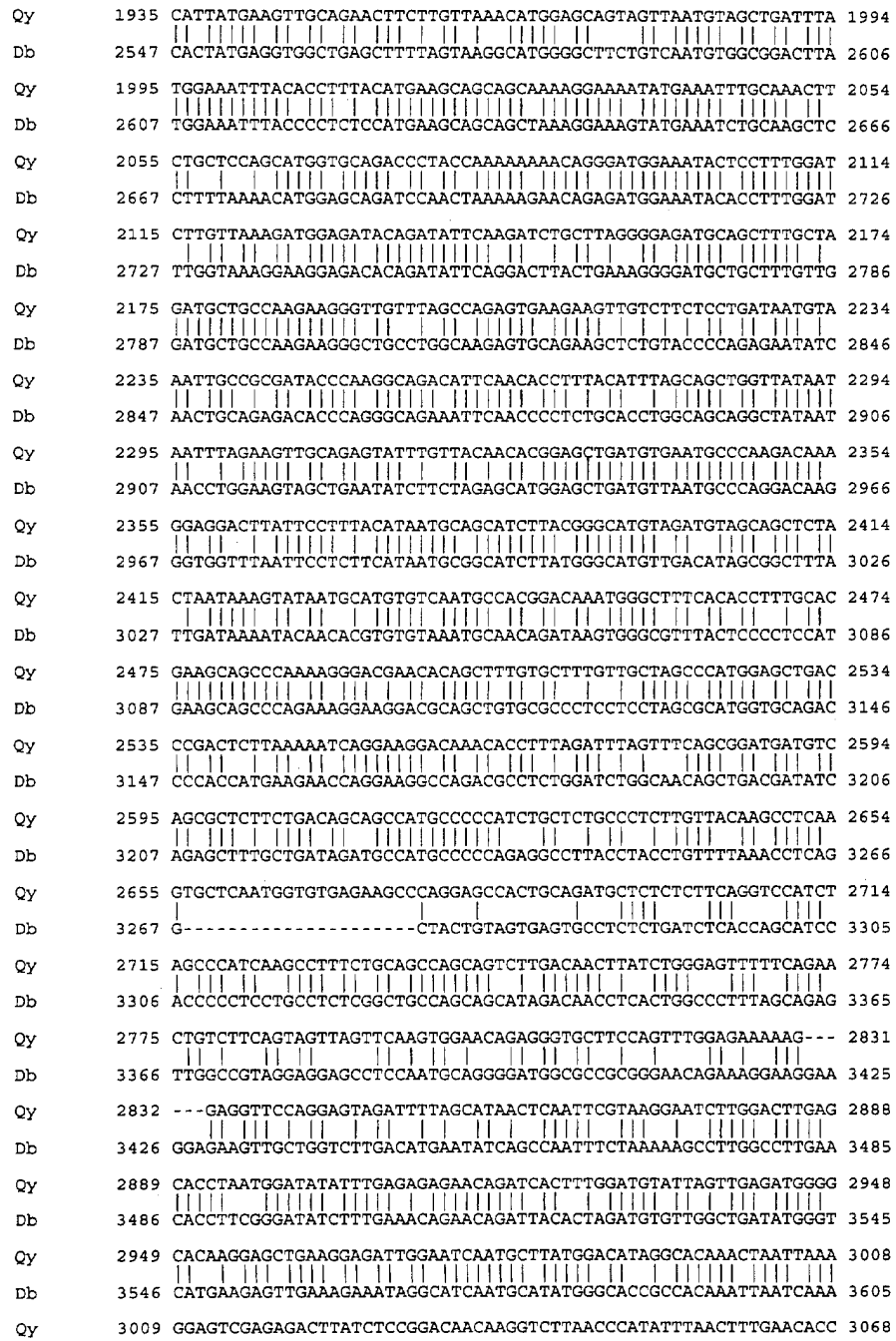
Qy      1664 ----------------------------CCTCCAAGAGGGTATCTCATTAGGTAATTCA 1694
Db      2247 GCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACAGAGTACACCTATACGTACTTCT 2306

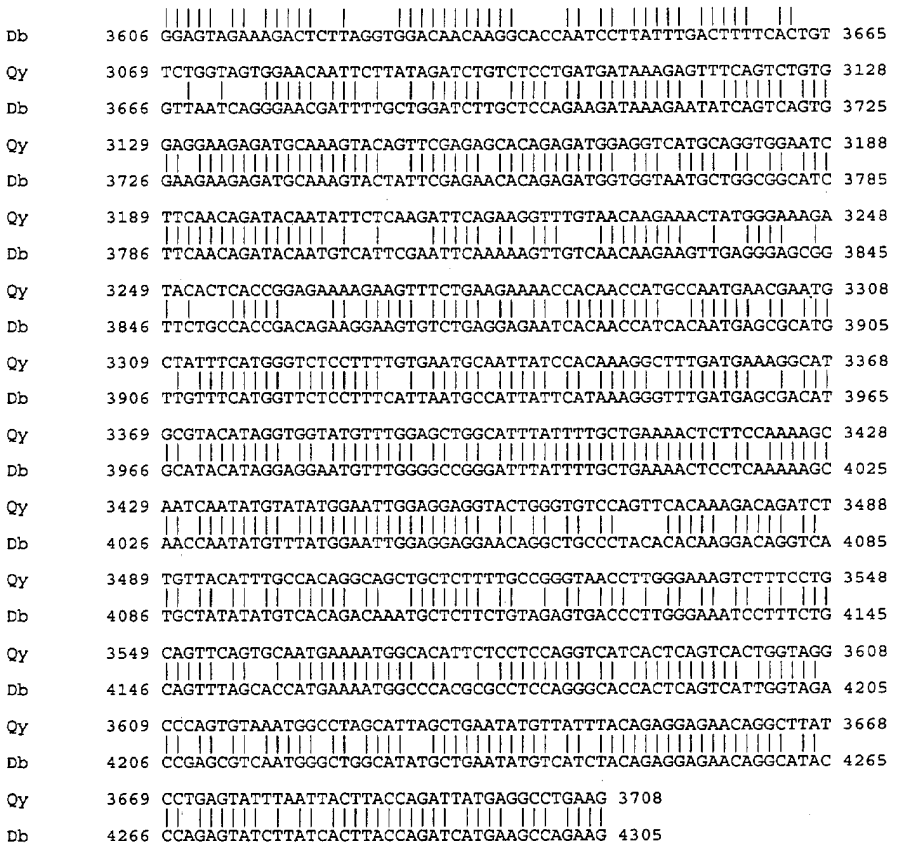
Qy      1695 GAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAA 1754
Db      2307 GATGTTGATTATCGACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAGCAA 2366

Qy      1755 CTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTT 1814
Db      2367 CTTTGCAGCTCTCAAAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCCTTA 2426

Qy      1815 CATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCT 1874
Db      2427 CACTTCGCAGCAGGCTACAACCGCGTGTCTGTTGTAGAGTACCTGCTACACCACGGTGCC 2486

Qy      1875 GATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGA 1934
Db      2487 GATGTCCATGCCAAAGACAAGGGTGGCTTGGTGCCCCTTCATAATGCCTGTTCATATGGA 2546
```

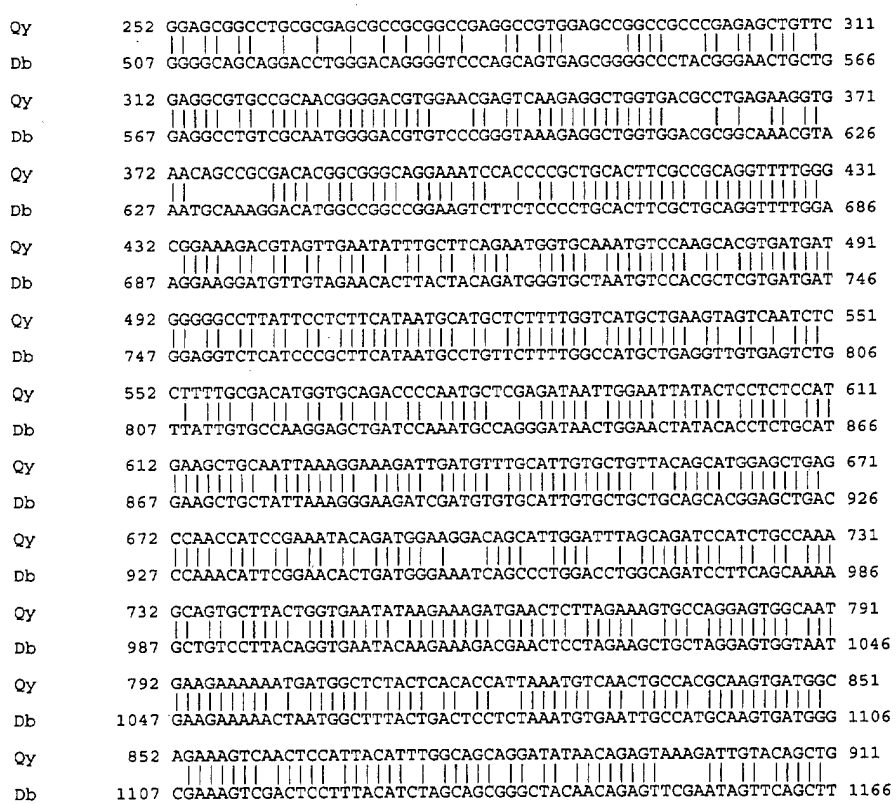


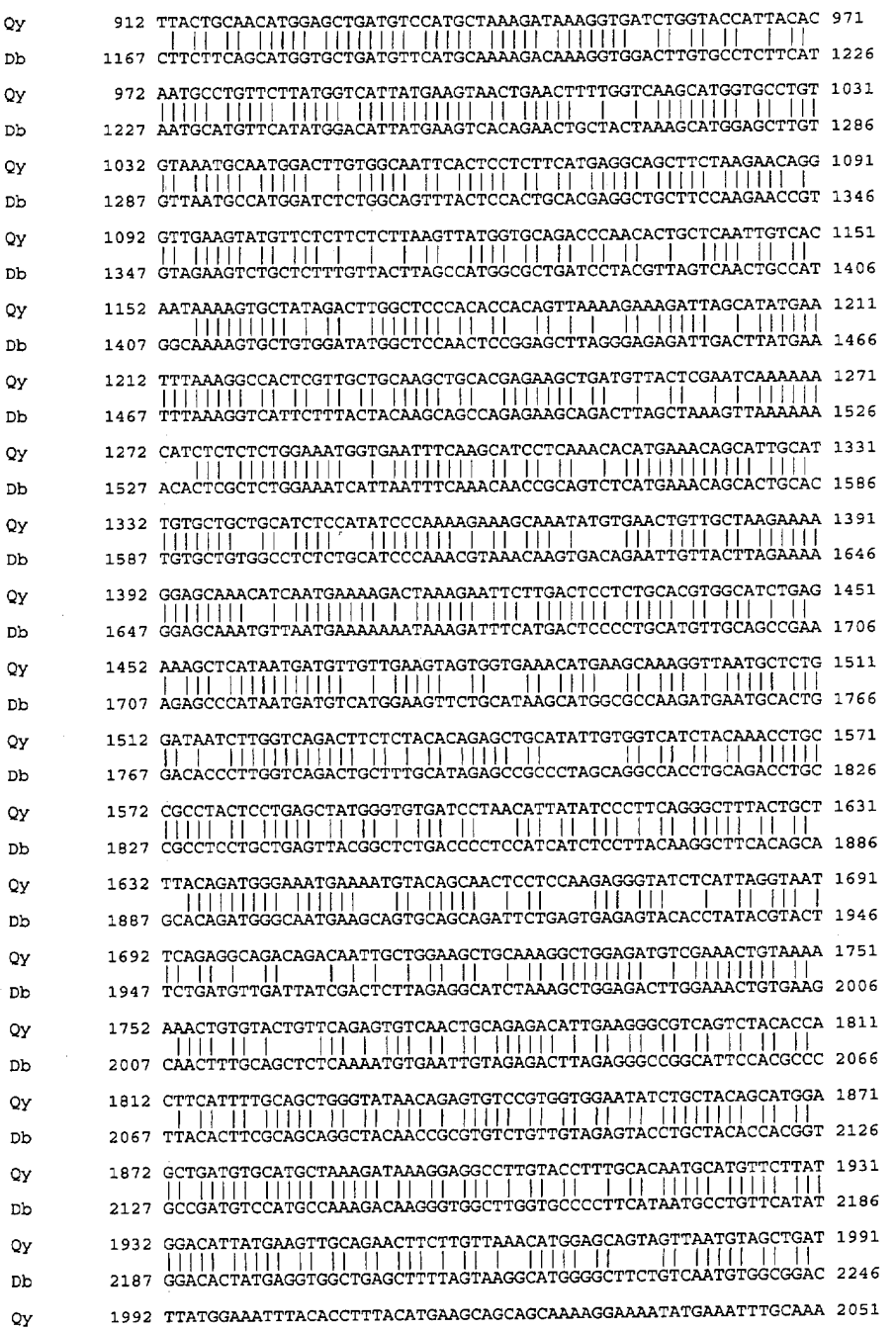


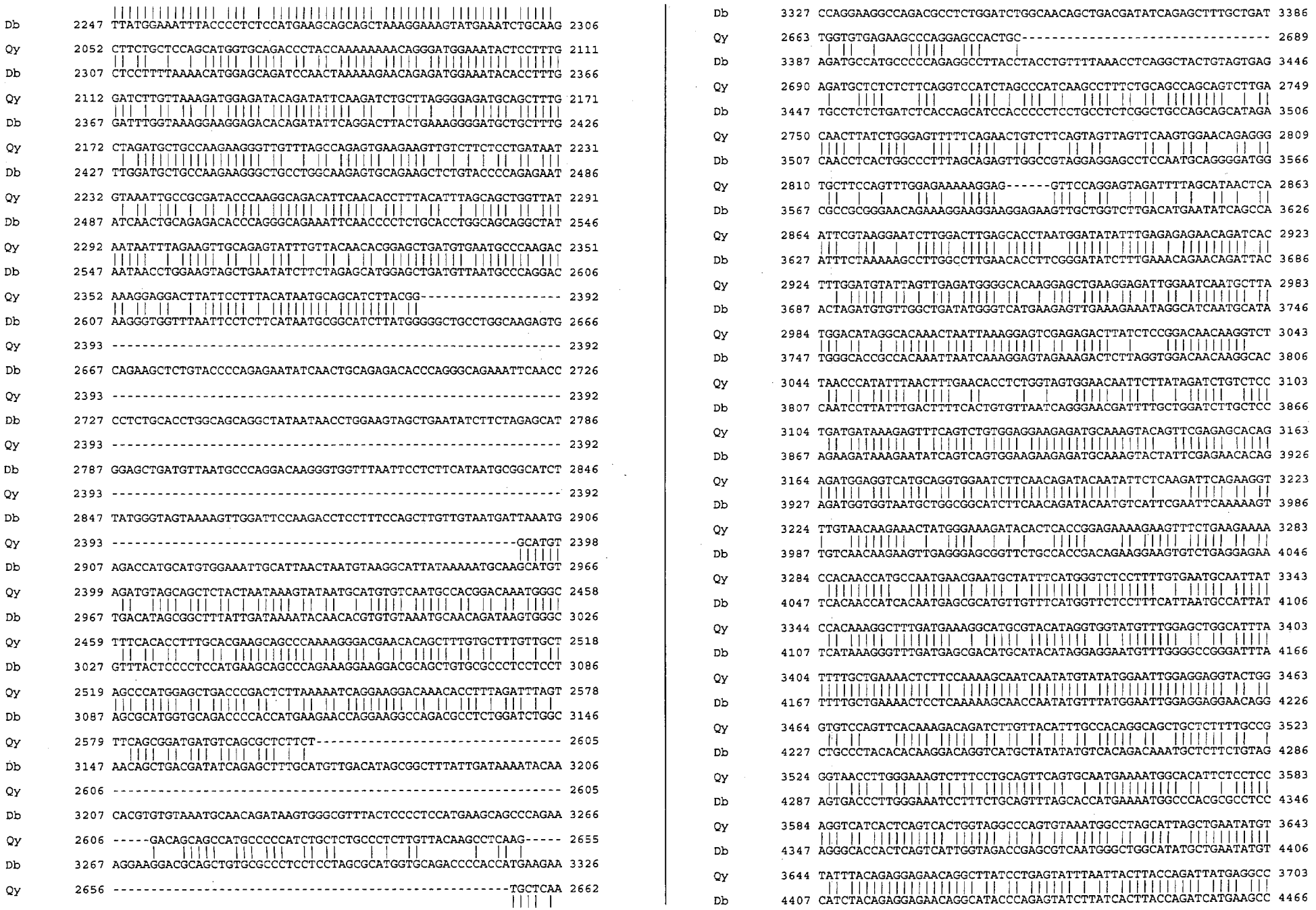
```
RESULT 22
US-09-841-835-9
; Sequence 9, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
```

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4657 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2855
US-09-841-835-9

  Query Match             35.9%;  Score 1369.4;  DB 9;  Length 4657;
  Best Local Similarity   63.1%;  Pred. No. 0;
  Matches 2501;  Conservative    0;  Mismatches  956;  Indels  508;  Gaps    5;
```







```
Qy       3704 TGAAG 3708
              ||||
Db       4467 AGAAG 4471
```

RESULT 23
US-10-276-774-340
; Sequence 340, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 340
; LENGTH: 1466
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-340

Query Match 32.5%; Score 1240.4; DB 12; Length 1466;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1247; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy       1042 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 60

Qy       1102 GTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 1161
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 GTTCTCTTCTCTTAAGTTATGGGGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 120

Qy       1162 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 1221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 180

Qy       1222 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 1281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 240

Qy       1282 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 1341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 300

Qy       1342 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 1401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 360

Qy       1402 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 1461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 420

Qy       1462 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 1521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 480

Qy       1522 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 1581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 540

Qy       1582 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 1641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 600

Qy       1642 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 1701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 660

Qy       1702 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTA 1761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTA 720

Qy       1762 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 1821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 780

Qy       1822 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 1881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 840

Qy       1882 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 1941
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 900

Qy       1942 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 2001
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 960

Qy       2002 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 2061
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 1020

Qy       2062 AGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 2121
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 1080

Qy       2122 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 2181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 1140

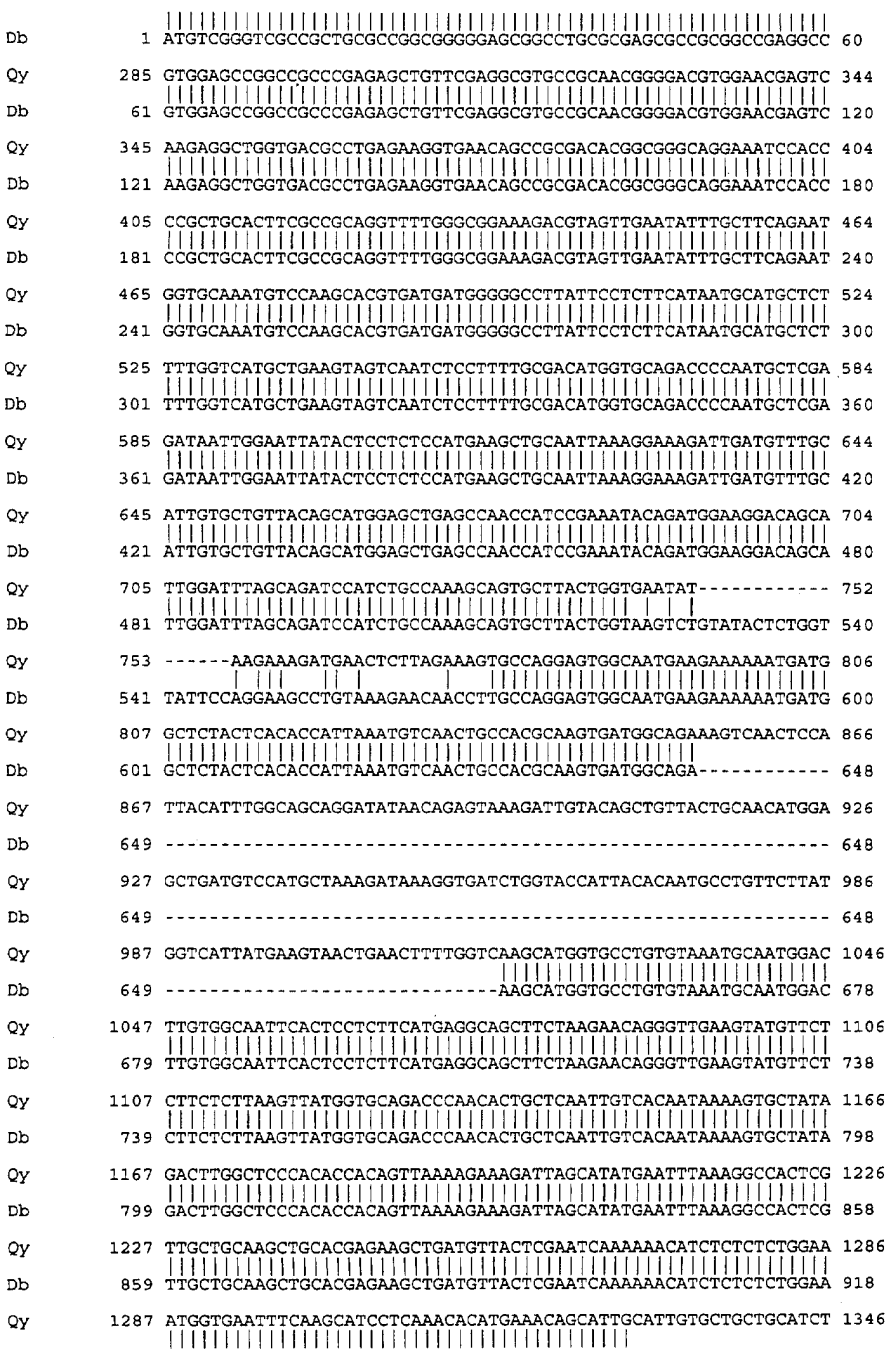
Qy       2182 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 2241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 1200

Qy       2242 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2299
              |||||||||||||||||||||||||||||||||||||||||| | | | || ||
Db        1201 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGGTAAGTGAATGAAGTT 1258
```

RESULT 24
US-09-964-899-40
; Sequence 40, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; TITLE OF INVENTION: Alzheimer's Disease Using Drosophila Melanogaster
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-964-899-40

Query Match 31.3%; Score 1192.6; DB 9; Length 2409;
Best Local Similarity 67.4%; Pred. No. 1.9e-309;
Matches 2372; Conservative 0; Mismatches 19; Indels 1128; Gaps 4;

```
Qy        225 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 284
```


Db 1 ATGTCGGGTCGCCGCTGCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCC 60
Qy 285 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 344
Db 61 GTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTC 120
Qy 345 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 404
Db 121 AAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACC 180
Qy 405 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 464
Db 181 CCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAAT 240
Qy 465 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 524
Db 241 GGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCT 300
Qy 525 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 584
Db 301 TTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGA 360
Qy 585 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 644
Db 361 GATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGC 420
Qy 645 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 704
Db 421 ATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCA 480
Qy 705 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATAT---------- 752
Db 481 TTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTAAGTCTGTATACTCTGGT 540
Qy 753 ------AAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATG 806
Db 541 TATTCCAGGAAGCCTGTAAAGAACAACCTTGCCAGGAGTGGCAATGAAGAAAAAATGATG 600
Qy 807 GCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCA 866
Db 601 GCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGA---------- 648
Qy 867 TTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGA 926
Db 649 ---------- 648
Qy 927 GCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTAT 986
Db 649 ---------- 648
Qy 987 GGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGAC 1046
Db 649 ----------AAGCATGGTGCCTGTGTAAATGCAATGGAC 678
Qy 1047 TTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCT 1106
Db 679 TTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCT 738
Qy 1107 CTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATA 1166
Db 739 CTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATA 798
Qy 1167 GACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCG 1226
Db 799 GACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCG 858
Qy 1227 TTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAA 1286
Db 859 TTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAA 918
Qy 1287 ATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCT 1346
Db 919 ATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTG---------- 960
Qy 1347 CCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAAT 1406
Db 961 ---------- 960
Qy 1407 GAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGAT 1466
Db 961 ---------- 960
Qy 1467 GTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAG 1526
Db 961 ---------- 960
Qy 1527 ACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGC 1586
Db 961 ---------- 960
Qy 1587 TATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAAT 1646
Db 961 ---------- 960
Qy 1647 GAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGA 1706
Db 961 ---------- 960
Qy 1707 CAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTT 1766
Db 961 ----------AAACTGTGTACTGTT 975
Qy 1767 CAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCT 1826
Db 976 CAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCT 1035
Qy 1827 GGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCT 1886
Db 1036 GGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCT 1095
Qy 1887 AAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTT 1946
Db 1096 AAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTT 1155
Qy 1947 GCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACA 2006
Db 1156 GCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACA 1215
Qy 2007 CCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCAT 2066
Db 1216 CCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCAT 1275
Qy 2067 GGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGAT 2126
Db 1276 GGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGAT 1335
Qy 2127 GGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAG 2186
Db 1336 GGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAG 1395
Qy 2187 AAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGAT 2246
Db 1396 AAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGAT 1455
Qy 2247 ACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTT 2306
Db 1456 ACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTT 1515
Qy 2307 GCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATT 2366
Db 1516 GCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATT 1575
Qy 2367 CCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTAT 2426
Db 1576 CCTTTACATAATGCAGCATCTTACGG---------- 1601

```
Qy       2427 AATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAA 2486
Db       1602 ------------------------------------------------------------ 1601
Qy       2487 AAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAA 2546
Db       1602 ------------------------------------------------------------ 1601
Qy       2547 AATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTG 2606
Db       1602 ------------------------------------------------------------ 1601
Qy       2607 ACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGT 2666
Db       1602 ------------------------------------------------------------ 1601
Qy       2667 GTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGC 2726
Db       1602 ------------------------------------------------------------ 1601
Qy       2727 CTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTA 2786
Db       1602 ------------------------------------------------------------ 1601
Qy       2787 GTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTA 2846
Db       1602 ------------------------------------------------------------ 1601
Qy       2847 GATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTT 2906
Db       1602 ------------------------------------------------------------ 1601
Qy       2907 GAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAG 2966
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1602 -----------GATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAG 1650
Qy       2967 ATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATC 3026
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 ATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATC 1710
Qy       3027 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 3086
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1711 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 1770
Qy       3087 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 3146
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1771 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 1830
Qy       3147 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 3206
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1831 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 1890
Qy       3207 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 3266
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 1950
Qy       3267 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 3326
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 2010
Qy       3327 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATG 3386
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATG 2070
Qy       3387 TTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGA 3446
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2071 TTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGA 2130
Qy       3447 ATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGG 3506
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2131 ATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGG 2190
```

```
Qy       3507 CAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAA 3566
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2191 CAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAA 2250
Qy       3567 ATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTA 3626
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2251 ATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTA 2310
Qy       3627 GCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACT 3686
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2311 GCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACT 2370
Qy       3687 TACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAA 3725
                |||||||||||||||||||||||||||||||||||||||
Db       2371 TACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAA 2409
```

RESULT 25
US-10-199-937-177
; Sequence 177, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 177
; LENGTH: 3308
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Parpla-Tank2b
; OTHER INFORMATION: Fusion
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3297)
US-10-199-937-177

Query Match 30.6%; Score 1168.4; DB 14; Length 3308;
Best Local Similarity 99.9%; Pred. No. 7.7e-303;
Matches 1169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2553 GAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCA 2612
                ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       2101 GAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCA 2160
Qy       2613 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2672
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2220
Qy       2673 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2732
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2280
Qy       2733 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2792
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2340
Qy       2793 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2852
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2400
Qy       2853 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2912
```

```
Db       2401 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2460

Qy       2913 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2972
Db       2461 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2520

Qy       2973 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 3032
Db       2521 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 2580

Qy       3033 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 3092
Db       2581 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 2640

Qy       3093 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 3152
Db       2641 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 2700

Qy       3153 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 3212
Db       2701 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 2760

Qy       3213 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 3272
Db       2761 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 2820

Qy       3273 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 3332
Db       2821 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 2880

Qy       3333 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 3392
Db       2881 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 2940

Qy       3393 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3452
Db       2941 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3000

Qy       3453 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3512
Db       3001 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3060

Qy       3513 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3572
Db       3061 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3120

Qy       3573 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3632
Db       3121 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3180

Qy       3633 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3692
Db       3181 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3240

Qy       3693 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
Db       3241 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3270
```

RESULT 26
US-10-199-937-159
; Sequence 159, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 159
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1094)
US-10-199-937-159

Query Match 28.4%; Score 1084.8; DB 14; Length 1103;
Best Local Similarity 99.8%; Pred. No. 1.1e-280;
Matches 1086; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1056 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 1115
Db          9 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 68

Qy       1116 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 1175
Db         69 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 128

Qy       1176 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 1235
Db        129 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 188

Qy       1236 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 1295
Db        189 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 248

Qy       1296 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 1355
Db        249 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 308

Qy       1356 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 1415
Db        309 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 368

Qy       1416 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 1475
Db        369 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 428

Qy       1476 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 1535
Db        429 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 488

Qy       1536 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 1595
Db        489 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 548

Qy       1596 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 1655
Db        549 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 608

Qy       1656 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 1715
Db        609 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 668

Qy       1716 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTC 1775
Db        669 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTC 728

Qy       1776 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 1835
Db        729 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 788

Qy       1836 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 1895
```

```
Db        789 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 848

Qy       1896 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 1955

Db        849 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 908

Qy       1956 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 2015

Db        909 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 968

Qy       2016 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 2075

Db        969 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 1028

Qy       2076 CCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 2135

Db       1029 CCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 1088

Qy       2136 GATATTCA 2143

Db       1089 GATATTTA 1096


RESULT 27
US-10-199-937-138
; Sequence 138, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 138
;   LENGTH: 5482
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (474)..(4016)
US-10-199-937-138

  Query Match             28.0%;  Score 1066.8;  DB 14;  Length 5482;
  Best Local Similarity   58.5%;  Pred. No. 2.4e-275;
  Matches 2037;  Conservative    0;  Mismatches 1332;  Indels  111;  Gaps    6;

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360

Db        547 GGGAGCTCTCCGAGGCCTGCAAAACGGGCGAGATCGCCAAGGTGAAGAAGCTAATAACGC 606

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420

Db        607 CTCAGACCGTGAACGCCAGGGATACGGCGGGACGCAAATCCACACCATTGCATTTCGCAG 666

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480

Db        667 CGGGTTATGGACGCCGGGAAGTGGTTGAATTCCTGCTGAACAGCGGCGCCTCCATACAGG 726

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540

Db        727 CGTGTGACGAGGGTGGGCTGCACCCGCTGCACAACTGTTGCTCCTTTGGCCACGCCGAGG 786

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 600

Db        787 TAGTTCGATTGTTGCTGAAGGCAGGTGCCAGTCCAAACACCACCGACAACTGGAACTACA 846

Qy        601 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 660

Db        847 CGCCATTGCACGAGGCGGCCAGCAAGGGCAAGGTGGATGTGTGCCTGGCTCTGTTGCAGC 906

Qy        661 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 720

Db        907 ATGGCGCAAACCATACGATCCGCAACTCGGAGCAGAAGACACCACTGGAGCTGGCGGACG 966

Qy        721 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 780

Db        967 AGGCGACGCGTCCCGTATTGACCGGCGAATATCGAAAGGATGAGCTGCTTGAAGCCGCAC 1026

Qy        781 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 840

Db       1027 GCTCGGGGGCCGAGGATCGCCTGCTGGCCCTACTCACGCCACTCAATGTCAACTGTCATG 1086

Qy        841 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 900

Db       1087 CCAGCGATGGACGACGCTCAACGCCGCTCCATCTGGCAGCGGGCTACAATCGGATCGGCA 1146

Qy        901 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 960

Db       1147 TCGTGGAAATTCTGCTGGCCAACGGAGCGGATGTACATGCTAAGGACAAGGGCGGTCTGG 1206

Qy        961 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1020

Db       1207 TGCCGCTGCACAATGCCTGCTCCTACGGACACTTCGATGTGACCAAGCTGCTTATCCAGG 1266

Qy       1021 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1080

Db       1267 CGGGCGCCAATGTCAACGCCAACGATCTGTGGGCCTTTACGCCGCTCCACGAGGCCGCCT 1326

Qy       1081 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1140

Db       1327 CCAAAAGTCGCGTCGAGGTCTGCAGCCTGCTGCTCAGTCGTGGAGCGGATCCCACCCTCC 1386

Qy       1141 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1200

Db       1387 TAAACTGCCACAGCAAGTCGGCCATCGATGCGGCGCCCACCAGGGAGCTGAGAGAGCGGA 1446

Qy       1201 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1260

Db       1447 TTGCCTTTGAATACAAGGGTCACTGCCTGCTGGACGCCTGTCGAAAGTGTGATGTGTCCC 1506

Qy       1261 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1320

Db       1507 GTGCCAAGAAGCTGGTATGCGCAGAGATTGTTAACTTCGTGCATCCATATACAGGAGACA 1566

Qy       1321 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1380

Db       1567 CTCCGCTCCACCTGGCCGTTGTCAGTCCGGATGGGAAGCGCAAGCAGCTGATGGAACTGC 1626

Qy       1381 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1440

Db       1627 TGACCAGAAAGGGATCCTTGCTGAACGAGAAAAACAAGGCTTTCCTCACGCCCCTGCATT 1686

Qy       1441 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1500

Db       1687 TGGCTGCCGAGCTGCTTCACTACGATGCCATGGAGGTGCTGCTAAAGCAGGGCGCCAAGG 1746

Qy       1501 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1560

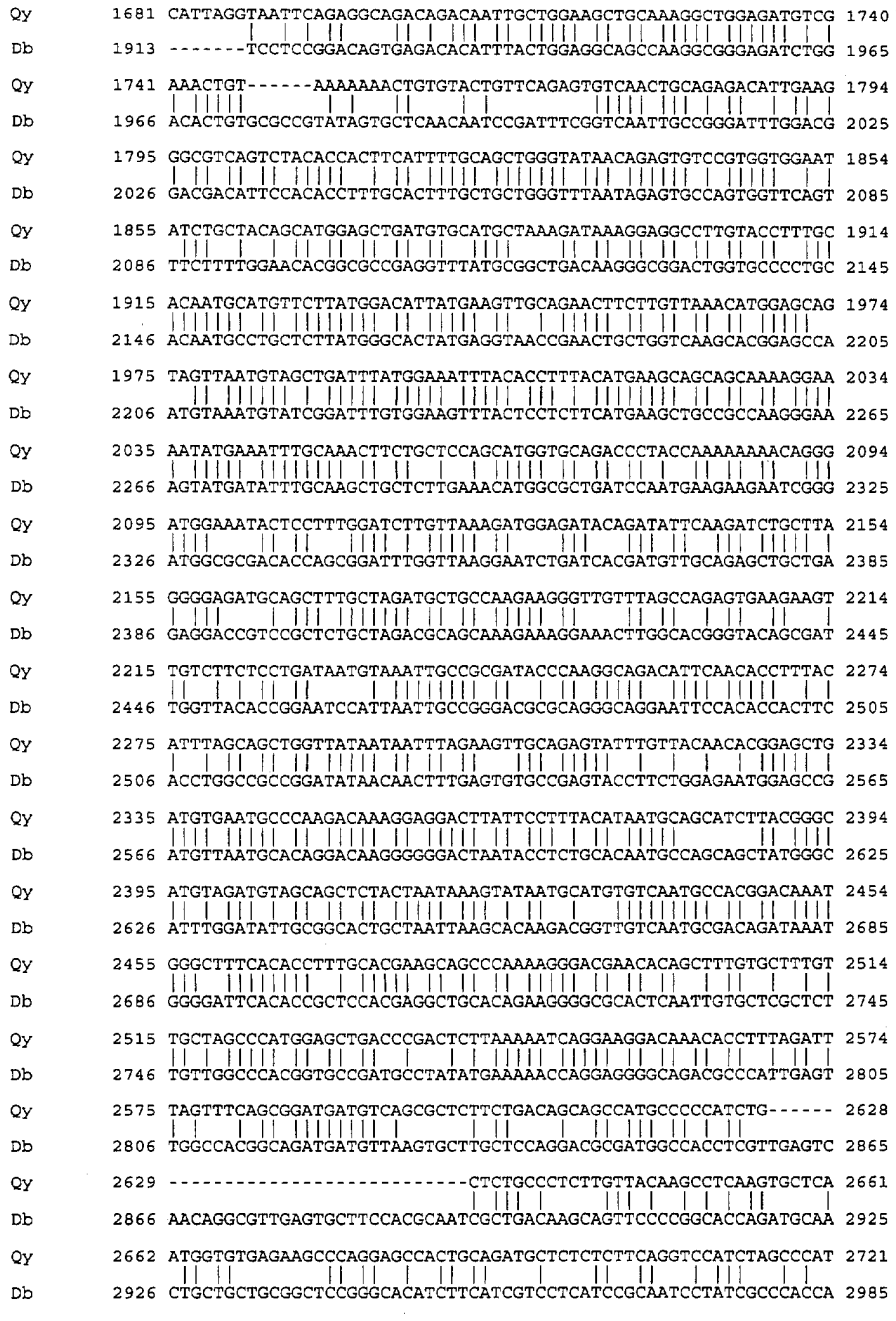
Db       1747 TTAATGCATTGGACAGTCTTGGACAAACGCCACTGCATCGGTGCGC------CCGTGATG 1800

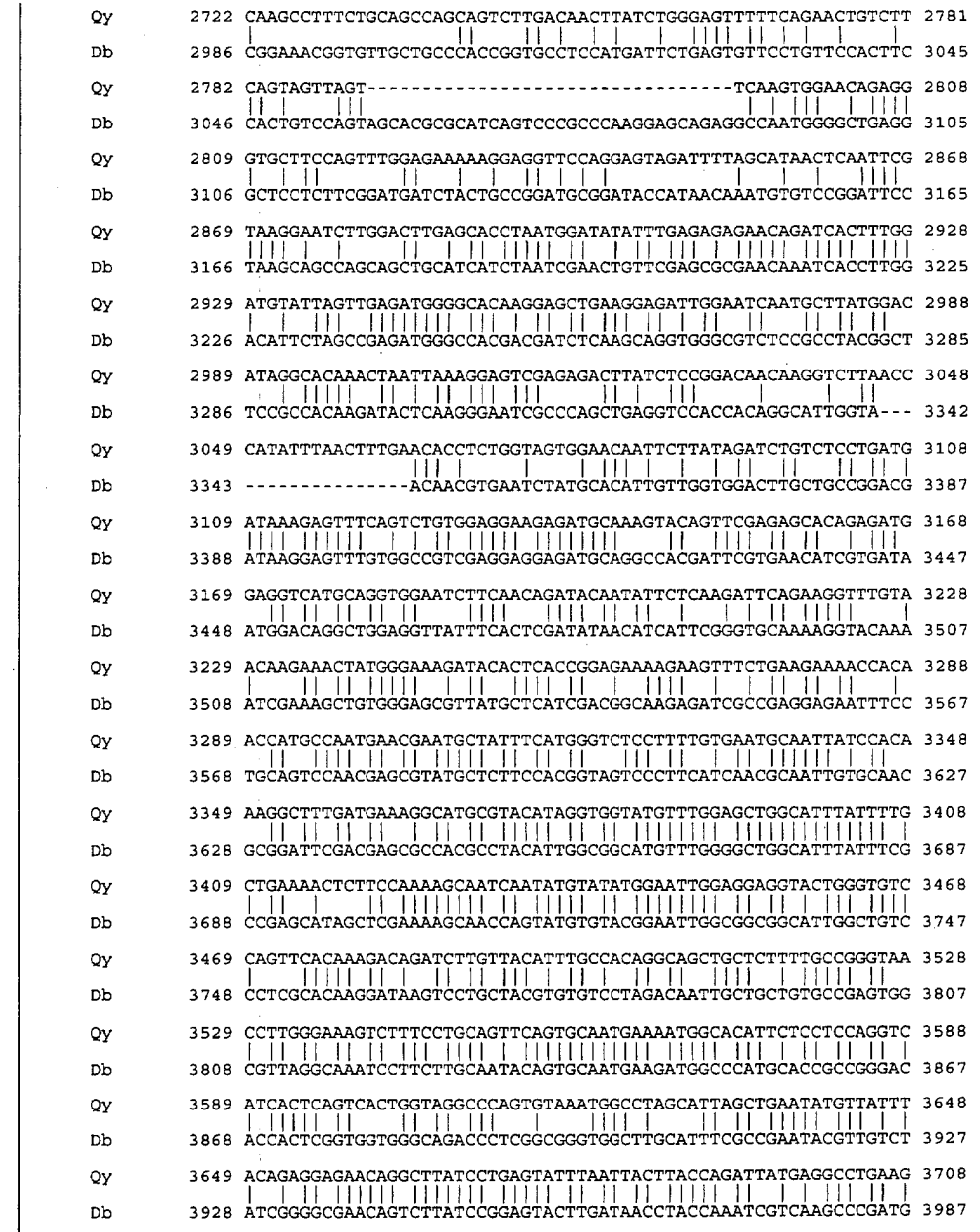
Qy       1561 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1620

Db       1801 AGCAAGCGGTGCGACTGCTGCTCTCGTACGCAGCGGACACGAATATCGTTTCCCTTGAGG 1860

Qy       1621 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1680

Db       1861 GACTTACGGCCGCTCAATTGGCCTCGGACAGCGTGCTGAAGCTGCTCAAGAA-------- 1912
```





RESULT 28
US-10-252-157-341
; Sequence 341, Application US/10252157
; Publication No. US20030190640A1

```
; GENERAL INFORMATION:
;  APPLICANT: Faris, Mary
;  APPLICANT:  Pearson, Cecelia I.
;  TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
;  FILE REFERENCE: PA-0027-1 US
;  CURRENT APPLICATION NUMBER: US/10/252,157
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,048
;  PRIOR FILING DATE: 2001-05-31
;  NUMBER OF SEQ ID NOS: 501
;  SOFTWARE: PERL Program
; SEQ ID NO 341
;   LENGTH: 797
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030190640A1 405041.1
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 257
;   OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-341

  Query Match            20.8%;  Score 794.4;  DB 14;  Length 797;
  Best Local Similarity  99.7%;  Pred. No. 1.2e-202;
  Matches  795;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1363 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 60

Qy       1423 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 120

Qy       1483 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 180

Qy       1543 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 240

Qy       1603 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 1662
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACATTATATCCCTTCANGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 300

Qy       1663 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 360

Qy       1723 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 420

Qy       1783 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 480

Qy       1843 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 540

Qy       1903 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 600

Qy       1963 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 660

Qy       2023 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 720

Qy       2083 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 780

Qy       2143 AAGATCTGCTTAGGGGA 2159
             ||||||||||||||||
Db        781 AAGATCTGCTTAGGGGA 797


RESULT 29
US-10-199-937-45
; Sequence 45, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
;   LENGTH: 582
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(480)
US-10-199-937-45

  Query Match            15.0%;  Score 574;  DB 14;  Length 582;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-143;
  Matches  574;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3243 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 60

Qy       3303 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 120

Qy       3363 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 180

Qy       3423 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 240

Qy       3483 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 300

Qy       3543 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 360

Qy       3603 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 420

Qy       3663 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 480

Qy       3723 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 540

Qy       3783 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
             ||||||||||||||||||||||||||||||||||
Db        541 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAA 574


RESULT 30
US-10-199-937-104
; Sequence 104, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 104
;   LENGTH: 647
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(646)
US-10-199-937-104

  Query Match             14.8%;  Score 563.6;  DB 14;  Length 647;
  Best Local Similarity   99.1%;  Pred. No. 1.3e-140;
  Matches  577;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 CGCGCTGCTCCGCCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 60
              ||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CGCGCTGCCTCCGCCGCCGCGGGGCAGCCGGGGGGCAGGGAGCCCAGCGAGGGGCGCGCG 126

Qy         61 TGGGCGCGGCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 120
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 TGGGCGCGG-CCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGG 185

Qy        121 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 CCCTGAGCGCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGC 245

Qy        181 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TCCGGTTGCTGGCGCTGTTGCTGGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCT 305

Qy        241 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GCGCCGGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCC 365

Qy        301 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GAGAGCTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGC 425

Qy        361 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        426 CTGAGAAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCG 485

Qy        421 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG 545

Qy        481 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 605

Qy        541 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTC 582
              |||||||||||||||||||||||||||||||||||||||||||
Db        606 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTC 647


RESULT 31
US-09-972-115A-10/c
; Sequence 10, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 520
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-10

  Query Match             13.6%;  Score 518.4;  DB 10;  Length 520;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-128;
  Matches  519;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        677 CATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 CATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGT 461

Qy        737 GCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGA 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGA 401

Qy        797 AAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAA 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 AAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAA 341

Qy        857 GTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACT 916
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACT 281

Qy        917 GCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGC 976
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGC 221

Qy        977 CTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAA 1036
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAA 161

Qy       1037 TGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGA 1096
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 TGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGA 101

Qy       1097 AGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAA 1156
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAA 41

Qy       1157 AAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAA 1196
                |||||||||||||||||||||||||||||||||||| ||
Db         40 AAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAAAA 1
```

RESULT 32
US-09-833-381-841
; Sequence 841, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 841
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1069)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-841

Query Match 12.3%; Score 467.6; DB 9; Length 1069;
Best Local Similarity 98.8%; Pred. No. 1.3e-114;
Matches 492; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```
Qy       1662 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1721
                | | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CCCTTCCTAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 626

Qy       1722 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1781
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 686

Qy       1782 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1841
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 746

Qy       1842 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1901
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 806

Qy       1902 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1961
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 866

Qy       1962 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2021
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 926

Qy       2022 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2081
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACA 986

Qy       2082 AAAAAAAACA-GGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATAT 2140
                |||||||||| |||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        987 AAAAAAAACAGGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGAT-CAGATAT 1045

Qy       2141 TCAAGATCTGCTTAGGGG 2158
                ||||||||||||||||||
Db       1046 TCAAGATCTGCTTAGGGG 1063
```

RESULT 33
US-09-833-381-1153
; Sequence 1153, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1153
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1153

Query Match 12.1%; Score 461; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 4.1e-113;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        572 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 631
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 64

Qy        632 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 691
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 124

Qy        692 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 751
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 184

Qy        752 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 811
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 244

Qy        812 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 871
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 304

Qy        872 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 931
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 364

Qy        932 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 991
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 424

Qy        992 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 1032
                |||||||||||||||||||||||||||||||||||||||
Db        425 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 465
```

RESULT 34
US-10-106-698-2223
; Sequence 2223, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03

```
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2223
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (136)..(136)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2223

  Query Match           12.0%;  Score 459;  DB 14;  Length 460;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-112;
  Matches  459;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2772 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 2831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 60

Qy       2832 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 2891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 120

Qy       2892 CTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 2951
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 180

Qy       2952 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 3011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 240

Qy       3012 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 3071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 300

Qy       3072 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 3131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 360

Qy       3132 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 3191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 420

Qy       3192 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 3231
             |||||||||||||||||||||||||||||||||||||||
Db        421 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 460


RESULT 35
US-10-199-937-5
; Sequence 5, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (136)
;   OTHER INFORMATION: n= a, c, g, or t
US-10-199-937-5

  Query Match           12.0%;  Score 459;  DB 14;  Length 460;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-112;
  Matches  459;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2772 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 2831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 60

Qy       2832 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 2891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 120

Qy       2892 CTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 2951
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 180

Qy       2952 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 3011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 240

Qy       3012 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 3071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 300

Qy       3072 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 3131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 360

Qy       3132 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 3191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 420

Qy       3192 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 3231
             |||||||||||||||||||||||||||||||||||||||
Db        421 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 460


RESULT 36
US-10-181-447A-11
; Sequence 11, Application US/10181447A
; Publication No. US20030180738A1
; GENERAL INFORMATION:
;  APPLICANT: The No. US20030180738A1tingham Trent University
;  TITLE OF INVENTION: Cancer Associated Genes and their Products
;  FILE REFERENCE: No. US20030180738A1tingham Trent Uni
;  CURRENT APPLICATION NUMBER: US/10/181,447A
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: PCT/GB/01/000188
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: GB0000993.6
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 402
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-181-447A-11

  Query Match           10.4%;  Score 396.8;  DB 14;  Length 402;
  Best Local Similarity   99.5%;  Pred. No. 7e-96;
  Matches  398;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        991 ATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGT 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGT 60

Qy       1051 GGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTC 1110
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTC 120

Qy        1111 TCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACT 1170
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACT 180

Qy        1171 TGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGC 1230
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 TGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGC 240

Qy        1231 TGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGG 1290
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGG 300

Qy        1291 TGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCAT 1350
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCAT 360

Qy        1351 ATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAA 1390
                |||||||||||||||||||||||||||||||||||  |||
Db         361 ATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAGAAAA 400


RESULT 37
US-09-972-115A-9
; Sequence 9, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-9

  Query Match             10.0%;  Score 380.4;  DB 10;  Length 384;
  Best Local Similarity   99.7%;  Pred. No. 1.8e-91;
  Matches  381;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1701 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 62

Qy        1761 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 122

Qy        1821 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 182

Qy        1881 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 242

Qy        1941 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 2000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 302

Qy        2001 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 362

Qy        2061 CAGCATGGTGCAGACCCTACCA 2082
                |||||||||||||||||||| |
Db         363 CAGCATGGTGCAGACCCTACAA 384


RESULT 38
US-09-918-995-17041
; Sequence 17041, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17041
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(441)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17041

  Query Match             9.6%;  Score 368;  DB 10;  Length 441;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-88;
  Matches  368;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3446 AATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAG 3505
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAG 60

Qy        3506 GCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAA 3565
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAA 120

Qy        3566 AATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCT 3625
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCT 180

Qy        3626 AGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAC 3685
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAC 240

Qy        3686 TTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTA 3745
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTA 300

Qy        3746 ATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3805
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 360

Qy        3806 AAAAAAAA 3813
                ||||||||
Db         361 AAAAAAAA 368


RESULT 39
US-10-199-937-41
; Sequence 41, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
```

```
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 1691
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(357)
US-10-199-937-41

  Query Match             9.4%;  Score 359.6;  DB 14;  Length 1691;
  Best Local Similarity   90.8%;  Pred. No. 2.1e-85;
  Matches  383;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 60

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 120

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 180

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 240

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 300

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGTA 360

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAA 3725
                  |||||  | |  | | | | |  |||  | |    |  |  |  || |
Db        361 ATGTAGTTTTATTTGTTCATCTTCAAAAATGCTAGGGAGGCATACTTTAACTTTTTATTA 420

Qy       3726 AT 3727
              ||
Db        421 AT 422


RESULT 40
US-10-199-937-43
; Sequence 43, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 1692
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(357)
US-10-199-937-43

  Query Match             9.4%;  Score 359.6;  DB 14;  Length 1692;
  Best Local Similarity   90.8%;  Pred. No. 2.1e-85;
  Matches  383;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       3306 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3365
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 60

Qy       3366 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3425
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 120

Qy       3426 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 180

Qy       3486 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 240

Qy       3546 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 300

Qy       3606 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGTA 360

Qy       3666 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAA 3725
                  |||||  | |  | | | | |  |||  | |    |  |  |  || |
Db        361 ATGTAGTTTTATTTGTTCATCTTCAAAAATGCTAGGGAGGCATACTTTAACTTTTTATTA 420

Qy       3726 AT 3727
              ||
Db        421 AT 422


RESULT 41
US-10-199-937-49
; Sequence 49, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
```

```
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
;    LENGTH: 356
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-199-937-49

  Query Match             9.3%;  Score 356;  DB 14;  Length 356;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-85;
  Matches  356;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3027 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 3086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 60

Qy       3087 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 3146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 120

Qy       3147 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 3206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 180

Qy       3207 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 3266
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 240

Qy       3267 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 3326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 300

Qy       3327 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGG 3382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGG 356


RESULT 42
US-10-199-937-7
; Sequence 7, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;    LENGTH: 564
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (203)
;    OTHER INFORMATION: n= a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (388)
;    OTHER INFORMATION: n= a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (441)
;    OTHER INFORMATION: n= a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (456)
;    OTHER INFORMATION: n= a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (538)
;    OTHER INFORMATION: n= a, c, g, or t
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (543)
;    OTHER INFORMATION: n= a, c, g, or t
US-10-199-937-7

  Query Match             9.3%;  Score 354.6;  DB 14;  Length 564;
  Best Local Similarity   93.4%;  Pred. No. 2e-84;
  Matches  369;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       3307 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 3366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 60

Qy       3367 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 3426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 120

Qy       3427 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 3486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 180

Qy       3487 CTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 3546
             ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db           181 CTTGTTACATTTGCCACAGGAGNCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 240

Qy       3547 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 3606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 300

Qy       3607 GGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTT 3666
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 GGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGTAA 360

Qy       3667 ATCCTGAGTATTTAATTACTTACCAGATTATGAGG 3701
                 |||||  | |  | | | | |   |||
Db           361 TGTAGTTTTATTTGTTCATCTTCAAAANTGCTAGG 395


RESULT 43
US-10-199-937-14
; Sequence 14, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;    LENGTH: 392
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (331)
```

```
;   OTHER INFORMATION: n= a, c, g, or t
US-10-199-937-14

  Query Match             9.1%;  Score 346;  DB 14;  Length 392;
  Best Local Similarity   99.7%;  Pred. No. 3.2e-82;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3307 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 3366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 60

Qy       3367 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 3426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 120

Qy       3427 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 3486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 180

Qy       3487 CTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 3546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 CTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 240

Qy       3547 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 3606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 300

Qy       3607 GGCCCAGTGTAAATGGCCTAGCATTAGCTG-AATATGTTATTTACAGAGGAGAACAGG 3663
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db           301 GGCCCAGTGTAAATGGCCTAGCATTAGCTGNAATATGTTATTTACAGAGGAGAACAGG 358


RESULT 44
US-10-199-937-12/c
; Sequence 12, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-199-937-12

  Query Match             8.9%;  Score 340.4;  DB 14;  Length 353;
  Best Local Similarity   99.4%;  Pred. No. 9.6e-81;
  Matches  352;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       3452 AGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCT 3511
             || |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db           353 AGAAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT-CATTTGCCACAGGCAGCT 295

Qy       3512 GCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 3571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           294 GCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 235

Qy       3572 ACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATT 3631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           234 ACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATT 175

Qy       3632 AGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCA 3691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           174 AGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCA 115

Qy       3692 GATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCA 3751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           114 GATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCA 55

Qy       3752 CTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3805
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            54 CTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 1


RESULT 45
US-10-040-739-1179
; Sequence 1179, Application US/10040739
; Publication No. US20020173635A1
;  GENERAL INFORMATION:
;       APPLICANT: Jacobs, Kenneth
;                  McCoy, John
;                  LaVallie, Edward
;                  Racie, Lisa
;                  Merberg, David
;                  Treacy, Maurice
;                  Spaulding, Vikki
;       TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;       NUMBER OF SEQUENCES: 1519
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy Disk
;            COMPUTER: IBM PC Compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/040,739
;            FILING DATE: 07-Jan-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/036,520
;            FILING DATE: 03-JUN-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Brown, Scott A.
;            REGISTRATION NUMBER: 32,724
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 498-8224
;            TELEFAX: (617) 876-5851
;  INFORMATION FOR SEQ ID NO: 1179:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 353 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1179:
US-10-040-739-1179

  Query Match             8.5%;  Score 323;  DB 13;  Length 353;
  Best Local Similarity   99.7%;  Pred. No. 4.6e-76;
  Matches  323;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3252 ACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTA 3311
             |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db            21 ACTCACCGGAGAAAAGAAGTTTCNGAAGAAAACCACAACCATGCCAATGAACGAATGCTA 80

Qy       3312 TTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCG 3371
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            81 TTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCG 140
```

```
Qy       3372 TACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAAT 3431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAAT 200

Qy       3432 CAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGT 3491
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGT 260

Qy       3492 TACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAG 3551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAG 320

Qy       3552 TTCAGTGCAATGAAAATGGCACAT 3575
             ||||||||||||||||||||||||
Db        321 TTCAGTGCAATGAAAATGGCACAT 344
```

Search completed: March 30, 2004, 17:12:06
Job time : 1342.71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 19:22:31 ; Search time 9107.17 Seconds
(without alignments)
12512.564 Million cell updates/sec

Title: US-09-843-159B-2
Perfect score: 3816
Sequence: 1 cgcgctgctccgcccgccgc.....cctttgctgaaaaaaaaaaa 3816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1663.4	43.6	1967	11	AK047094	AK047094 Mus muscu
	2	1609.8	42.2	3417	11	AK048860	AK048860 Mus muscu
	3	1436.2	37.6	3194	11	AK080913	AK080913 Mus muscu
	4	784.2	20.6	838	9	AU140145	AU140145 AU140145
	5	687.8	18.0	955	14	CB183123	CB183123 AGENCOURT
	6	672.4	17.6	675	14	CB128566	CB128566 K-EST0178
	7	672.2	17.6	1128	12	BM457025	BM457025 AGENCOURT
c	8	665.2	17.4	754	14	CB170279	CB170279 CEY603161
c	9	652	17.1	923	14	CF407574	CF407574 CH3#049_C
	10	649.2	17.0	659	14	R64714	R64714 EST54a09 WA
	11	643.8	16.9	691	12	BG391376	BG391376 602417447
	12	642.8	16.8	759	9	AL601027	AL601027 DKFZp313G
	13	634.4	16.6	950	13	BQ885764	BQ885764 AGENCOURT
	14	632.6	16.6	743	14	CF539235	CF539235 UI-M-GH0-
	15	622.8	16.3	638	13	BX474686	BX474686 DKFZp686J
	16	616.4	16.2	724	14	CF533694	CF533694 UI-M-FY0-
	17	614.2	16.1	747	14	CF727485	CF727485 UI-M-HB0-
	18	605	15.9	648	12	BG875327	BG875327 RC1-BT062
	19	596	15.6	703	14	CD803519	CD803519 UI-M-GV0-
c	20	593.2	15.5	622	10	BE220005	BE220005 hv65h03.x
c	21	584	15.3	616	10	BE467629	BE467629 hz67b11.x
	22	583.8	15.3	793	13	BU466115	BU466115 603371794
	23	578.8	15.2	689	14	CF739820	CF739820 UI-M-HD0-
	24	574	15.0	796	13	BU427328	BU427328 603234467
	25	572.8	15.0	742	14	CF735419	CF735419 UI-M-HB0-
	26	572.6	15.0	913	14	CB201923	CB201923 AGENCOURT
	27	570.4	14.9	956	10	BE910901	BE910901 601661978
	28	568	14.9	702	10	BB339554	BB339554 BB339554
	29	564.2	14.8	725	14	CF538288	CF538288 UI-M-GH0-
c	30	562.8	14.7	585	10	BE222311	BE222311 hu10d01.x
	31	561.2	14.7	681	14	CF728039	CF728039 UI-M-HB0-
	32	557.8	14.6	669	14	CF745978	CF745978 UI-M-GV0-
	33	557.6	14.6	659	10	BB646543	BB646543 BB646543
c	34	557.4	14.6	595	10	BE504175	BE504175 hv86c11.x
	35	552	14.5	640	14	CF530741	CF530741 UI-M-FY0-
	36	549	14.4	553	13	BQ267594	BQ267594 ij90c10.y
	37	545	14.3	691	14	CF729155	CF729155 UI-M-HD0-
	38	544	14.3	1099	12	BG245969	BG245969 602358786
c	39	538.6	14.1	562	10	BE504478	BE504478 hz57a02.x
	40	537.8	14.1	621	14	CF538447	CF538447 UI-M-GI0-
	41	537.6	14.1	703	14	CF734609	CF734609 UI-M-HB0-
	42	535.8	14.0	765	13	BU570530	BU570530 AGENCOURT
	43	535.6	14.0	679	14	CF726657	CF726657 UI-M-HB0-
	44	534.4	14.0	996	10	BF585334	BF585334 602102008
	45	533.2	14.0	769	14	CA317519	CA317519 UI-M-FW0-

ALIGNMENTS

RESULT 1
AK047094
LOCUS AK047094 1967 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930018O21 product:TANKYRASE-RELATED PROTEIN (FRAGMENT) homolog [Homo sapiens], full insert sequence.
ACCESSION AK047094
VERSION AK047094.1 GI:26338577
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1967)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES Location/Qualifiers
source 1. .1967
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:B930018O21"
/db_xref="MGI:2412634"
/db_xref="taxon:10090"
/clone="B930018O21"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
CDS 166. .>1965
/note="unnamed protein product; TANKYRASE-RELATED PROTEIN (FRAGMENT) homolog [Homo sapiens] (SPTR|Q9HAS4, evidence: FASTY, 94.6%ID, 51.4%length, match=1959)
putative"
/codon_start=1
/protein_id="BAC32960.1"
/db_xref="GI:26338578"
/translation="MSGRRCAGGGAACASAGAEAVEPSARELFEACRNGDVERVKRLV
TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
HAEVVNLLLQHGAHPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
LDLADPSAKAVLTGDYKEDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
AEYNRVKIVQLLLHHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
QFTPLHEAASKNRIEVCSLLLSYGADPTLLNCHNKSAIDLAPTAQLKERLSYEFKGHS
LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
TNEKTKEFLTPLHVASENAHNDVVEVVVKHEAKVNALDSLGQTSLHRAAHCGHLQTCR
LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGASLGHSEADRQLLEAAKAGDVETV
KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA"

ORIGIN

Query Match 43.6%; Score 1663.4; DB 11; Length 1967;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 1782; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

```
Qy      69 GCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCCTGAGC 128
           | |||||||| | ||||  || |||||||| ||||||  ||||||||||||||||||||
Db       7 GGCCATGGGACCGTGCCGACTCTGGTGACAGCGGGGAGCGGAGCGGCCCGGGCCCTGAGC 66

Qy     129 GCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCCGGTTG 188
           |||||| || |||||||||||||||||||||||||||||||||||||||||||| |  |||
Db      67 GCGTCTCCTGCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCCGGGCTCCTGCTGCCTTTG 126

Qy     189 CTGGCGCTGTTGCT---GGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCC 245
           ||||||||| ||||    ||| | |||||||||| |||||||||| || ||||||||||||
Db     127 CTGGCGCTGCTGCTGGCGGCGGCGGCGGCGGCCCGGATCATGTCTGGCCGCCGCTGCGCC 186

Qy     246 GGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAG 305
           |||||||| |||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db     187 GGCGGGGGCGCGGCCTGCGCGAGCGCCGGGGCCGAGGCCGTGGAGCCGTCCGCCCGCGAG 246

Qy     306 CTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAG 365
           |||||||||| |||||||||||||||||||| ||||||||||||||||||||||| |||
Db     247 CTGTTCGAGGCTTGCCGCAACGGGGACGTGGAGCGAGTCAAGAGGCTGGTGACGCCCGAG 306

Qy     366 AAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGT 425
           |||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db     307 AAGGTGAACAGCCGCGACACGGCGGGCAGGAAGTCCACCCCGCTCCACTTCGCCGCAGGT 366

Qy     426 TTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGT 485
           ||||| |||| || |||||||||| | |||||| |||||||||| ||||| |||||| |||
Db     367 TTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGCAAGCGCGT 426

Qy     486 GATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTC 545
           |||||||||| |||||||||||||||| |||||||||||||||||||||||||||||||||
Db     427 GATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTGAAGTAGTC 486

Qy     546 AATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCT 605
           ||||| ||||| | ||| |||||| |||| |||||||| ||||||||||||||||||||||
Db     487 AATCTTCTTTTACAACACGGTGCACACCCAAATGCTCGGGATAATTGGAATTATACTCCT 546

Qy     606 CTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGA 665
           |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||||
Db     547 CTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTACAGCATGGA 606

Qy     666 GCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCT 725
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     607 GCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGACCCATCT 666

Qy     726 GCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGT 785
           |||||||| || |||||||| ||||||||||  | ||||||||||||||||||||||||||
Db     667 GCCAAAGCTGTTCTGACTGGTGACTATAAGGAAGATGAACTTCTGGAAAGTGCCAGGAGT 726

Qy     786 GGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGT 845
           |||||||| |||||||||||||| || || |||||||| |||||||||||||| ||||||
Db     727 GGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCCATGCAAGT 786

Qy     846 GATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTA 905
           |||||||||||||||||||||||||||||||| ||| ||||||||||| ||||||||||||
Db     787 GATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGAATACAACAGAGTAAAAATTGTA 846
```

```
Qy        906 CAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCA 965
              || |||||||||| ||||| |||||||||||||||||||||||| ||||||||||||
Db        847 CAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATCTGGTACCA 906

Qy        966 TTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGT 1025
              |||||||||||| ||||| ||||||||||||||||||||||| |||||||||| |||
Db        907 CTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCACGGT 966

Qy       1026 GCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAG 1085
              |||||||| |||||||||| ||||||| |||||||||||||||||||||||||||||||
Db        967 GCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAGCTTCTAAG 1026

Qy       1086 AACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAAT 1145
              |||||| |||||||||||||||| ||| ||||||| || ||||| ||||| ||||| ||
Db       1027 AACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCAACCCTGCTGAAC 1086

Qy       1146 TGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCA 1205
              |||||||||||||||||||||| |||| ||||| ||||| | || |||||||||| |
Db       1087 TGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAAGATTATCG 1146

Qy       1206 TATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATC 1265
              |||||||||||||||| || ||||||||||||||||||||||||||  |||||
Db       1147 TATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTACAAGAATC 1206

Qy       1266 AAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCA 1325
              |||||||||| |||||||||||||||||||| |||||||| ||||||||||||||||
Db       1207 AAAAAACATCTTTCTCTGGAAATGGTGAATTTCAAACATCCTCAGACACATGAAACAGCA 1266

Qy       1326 TTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTA 1385
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||| ||||
Db       1267 CTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAGATATGTGAACTGCTGCTC 1326

Qy       1386 AGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCA 1445
              ||||| |||||||||| ||| ||||||||||||||||||||||||||||||||||||||
Db       1327 AGAAAGGGAGCAAACACCAACGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCT 1386

Qy       1446 TCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAAT 1505
              |||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1387 TCTGAGAACGCTCACAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAAT 1446

Qy       1506 GCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAA 1565
              ||| |||||| ||||| ||  |||||||||||||| | ||||||| || ||
Db       1447 GCTTTGGATAGTCTTGGACAGACGTCATTACACAGAGCTGCACACTGTGGTCACCTGCAG 1506

Qy       1566 ACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTT 1625
              |||||||||| |||||||||||||||||||||||| ||||| |||||||||||||| ||
Db       1507 ACCTGCCGCCTGCTCCTGAGCTATGGGTGTGATCCAAACATCATATCCCTTCAGGGTTTC 1566

Qy       1626 ACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTA 1685
              || || || || |||||||||||||||||| ||||| || |||||||||||    |||| ||
Db       1567 ACCGCCTTGCAAATGGGAAATGAAAATGTGCAGCAGCTGCTCCAAGAGGGCGCCTCACTA 1626

Qy       1686 GGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACT 1745
              ||| | |||||||| |||||||| |||||||||||||||||||||| || || ||||||
Db       1627 GGTCACTCAGAGGCGGACAGACAGCTGCTGGAAGCTGCAAAGGCTGGTGACGTTGAAACT 1686

Qy       1746 GTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCT 1805
              |||||||||| |||||||||||||||||||||||||||||||||||||| || |||||
Db       1687 GTAAAAAAACTCTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCAGTCA 1746

Qy       1806 ACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAG 1865
              || ||||| || |||||||||||||| |||||||||| ||||| |||||||||||| ||
Db       1747 ACCCCACTCCACTTTGCAGCTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAA 1806

Qy       1866 CATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGT 1925
              |||||||||||||| |||||||||| |||||||| |||||||||| ||||||||||||
Db       1807 CATGGAGCTGATGTTCATGCTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGCATGT 1866

Qy       1926 TCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTA 1985
              |||||||||| |||||||||||||||||||||||| || |||||||||||| |||||
Db       1867 TCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTC 1926

Qy       1986 GCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGC 2026
```

```
                  ||||| || ||||| || |||||||||||||||| |||||
Db       1927 GCTGACTTGTGGAAGTTCACACCTTTACATGAAGCTGCAGC 1967
```

RESULT 2
AK048860

LOCUS	AK048860 3417 bp mRNA linear HTC 20-SEP-2003
DEFINITION	Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched library, clone:C230076L23 product:TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE homolog [Homo sapiens], full insert sequence.
ACCESSION	AK048860
VERSION	AK048860.1 GI:26339607
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1
AUTHORS	Carninci,P. and Hayashizaki,Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 3417)
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers

source 1. .3417
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C230076L23"
/db_xref="MGI:2415999"
/db_xref="taxon:10090"
/clone="C230076L23"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"

CDS 423. .3353
/note="unnamed protein product; TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE homolog [Homo sapiens] (SPTR|O95271, evidence: FASTY, 95.9%ID, 80%length, match=3167)
putative"
/codon_start=1
/protein_id="BAC33475.1"
/db_xref="GI:26339608"
/translation="MALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADV
HAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSL
LLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLAL
EIINFKQPQSHETALHCAVASLHPKRKQVAELLLRKGANVNEKNKDFMTPLHVAAERA
HNDVMEVLHKHGAKMNALDSLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTA
AQMGNEAVQQILSESTPMRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHS
TPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASV
NVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLL
RGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEH
GADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQ
LCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVVSAS
VISPASTPSCLSAASSIDNLTGPLTDLAVGGASNAGDGAAGAERKEGEVAGLDMNIIQ
FLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQ
GTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIR
IQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGM
FGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFST
MKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTATAA
EQKT"

ORIGIN

Query Match 42.2%; Score 1609.8; DB 11; Length 3417;
Best Local Similarity 72.4%; Pred. No. 0;
Matches 2126; Conservative 0; Mismatches 782; Indels 27; Gaps 2;

```
Qy        780 AGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCAC 839
              ||||||||  ||||||||||||  | ||||||  ||||  ||  ||  |||||||  ||  |||||
Db        399 AGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCAT 458

Qy        840 GCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAG 899
              ||||||||||  |||||||  |||||  |||||  ||||||||||  ||  ||||||||     |
Db        459 GCAAGTGATGGACGAAAGTCTACTCCGTTACACCTGGCAGCAGGCTACAACAGAGTTCGG 518

Qy        900 ATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTG 959
              ||  ||  ||||||  |  ||  ||  |||||  ||||||||||  ||  |||||  |||||||   ||
Db        519 ATAGTTCAGCTGCTTCTACAGCATGGTGCTGATGTCCACGCGAAAGACAAAGGTGGACTT 578

Qy        960 GTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAG 1019
              |||||  |  ||  |||||  ||  ||  ||  ||  ||  |||||  ||  |  |||
Db        579 GTACCTCTTCATAATGCATGCTCATACGGACACTATGAGGTCACAGAACTGCTGCTAAAG 638

Qy       1020 CATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCT 1079
              |||||  ||  |||||  ||  ||  |||||  |||||||  |||||  ||  ||  |||||  |||
Db        639 CATGGAGCTTGTGTCAACGCTATGGATCTGTGGCAGTTCACACCGCTGCATGAAGCTGCT 698

Qy       1080 TCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTG 1139
              ||  ||||||  ||||  |||||  ||||||||  |  |  |||  |  |||||  ||  |||||  ||
Db        699 TCAAAGAACCGGGTAGAAGTCTGTTCTCTGTTACTTAGTCACGGTGCCGATCCAACCTTG 758

Qy       1140 CTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGA 1199
              |||||||  ||  |||||||||  |  |||  ||||||  ||  ||  |  |  ||||||
Db        759 GTCAATTGCCATGGCAAAAGTGCTGTCGACATGGCTCCAACTCCTGAGCTCCGGGAAAGA 818

Qy       1200 TTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACT 1259
              |  ||||||||||||||||||  ||  ||  |||||  ||  |||||||  ||  |  |
Db        819 CTGACATATGAATTTAAAGGCCATTCTTTACTCCAAGCCGCCAGAGAAGCAGACCTGGCC 878

Qy       1260 CGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAA 1319
              |  |  ||||||  ||  ||  |||||||  |  ||  ||||||||  ||  ||  |  ||||||
Db        879 AAAGTTAAAAAAACGCTTGCTTTGGAAATCATTAACTTCAAGCAGCCACAGTCTCATGAA 938

Qy       1320 ACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTG 1379
              ||  ||||||||  ||||||  ||  ||  |||||||  ||||||  |  ||  |||
Db        939 ACTGCATTGCACTGTGCTGTGGCCTCCTTGCATCCCAAACGAAAGCAGGTGGCAGAGCTG 998

Qy       1380 TTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCAC 1439
              ||  ||  ||  ||  |||||  |||  ||||  |||||||  ||  ||  |||  ||||  ||||||||
Db        999 TTACTGAGGAAGGGAGCCAACGTCAACGAAAAGAACAAGGATTTCATGACCCCTCTGCAT 1058

Qy       1440 GTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAG 1499
              ||  |||  ||||  |  ||||||||||||||  |  |||||  ||  ||  ||  |  |||||||
Db       1059 GTTGCAGCTGAAAGAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCACGGAGCAAAG 1118

Qy       1500 GTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCAT 1559
              |  |||||  |||||||  ||||||||||||||  ||  |  ||||||||||  |  ||  ||
Db       1119 ATGAATGCACTGGATAGTCTTGGTCAGACTGCTTTGCACAGAGCTGCCTTAGCGGGCCAC 1178

Qy       1560 CTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAG 1619
              ||  ||  ||||||||||||  ||  |||||  |||||  |  |||||||  |||  ||  |||||  ||
Db       1179 CTGCAGACCTGCCGCCTCCTGCTGAGTTATGGCTCTGATCCTTCCATCATCTCCCTGCAA 1238

Qy       1620 GGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATC 1679
              |||||  ||  ||  |||||||  |||||  ||  ||  ||  ||  ||  |||  |||
Db       1239 GGCTTCACGGCAGCGCAGATGGGGAATGAAGCAGTGCAACAGATTCTGAGCGAGAGTACT 1298

Qy       1680 TCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTC 1739
              |  |  |  |||  ||  |  |||  |  |  ||  ||  |  ||  |||||  ||  |
Db       1299 CCTATGCGCACTTCGGACGTGGACTATCGTCTCTTAGAGGCATCGAAAGCTGGTGACTTG 1358

Qy       1740 GAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGT 1799
              |||||||  ||  ||||  ||  |  ||||  |||  ||||||||||||  |  ||  ||  ||
Db       1359 GAAACTGTGAAGCAACTCTGCAGCCCGCAGAATGTGAACTGCAGAGACCTGGAGGGCCGG 1418

Qy       1800 CAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTG 1859
              ||  ||||||||  |  ||  ||  ||  ||  ||  ||||||  |  ||||  ||  ||  ||  ||
Db       1419 CACTCTACACCCTTACACTTCGCTGCAGGCTATAACCGTGTGTCTGTCGTTGAGTACCTT 1478

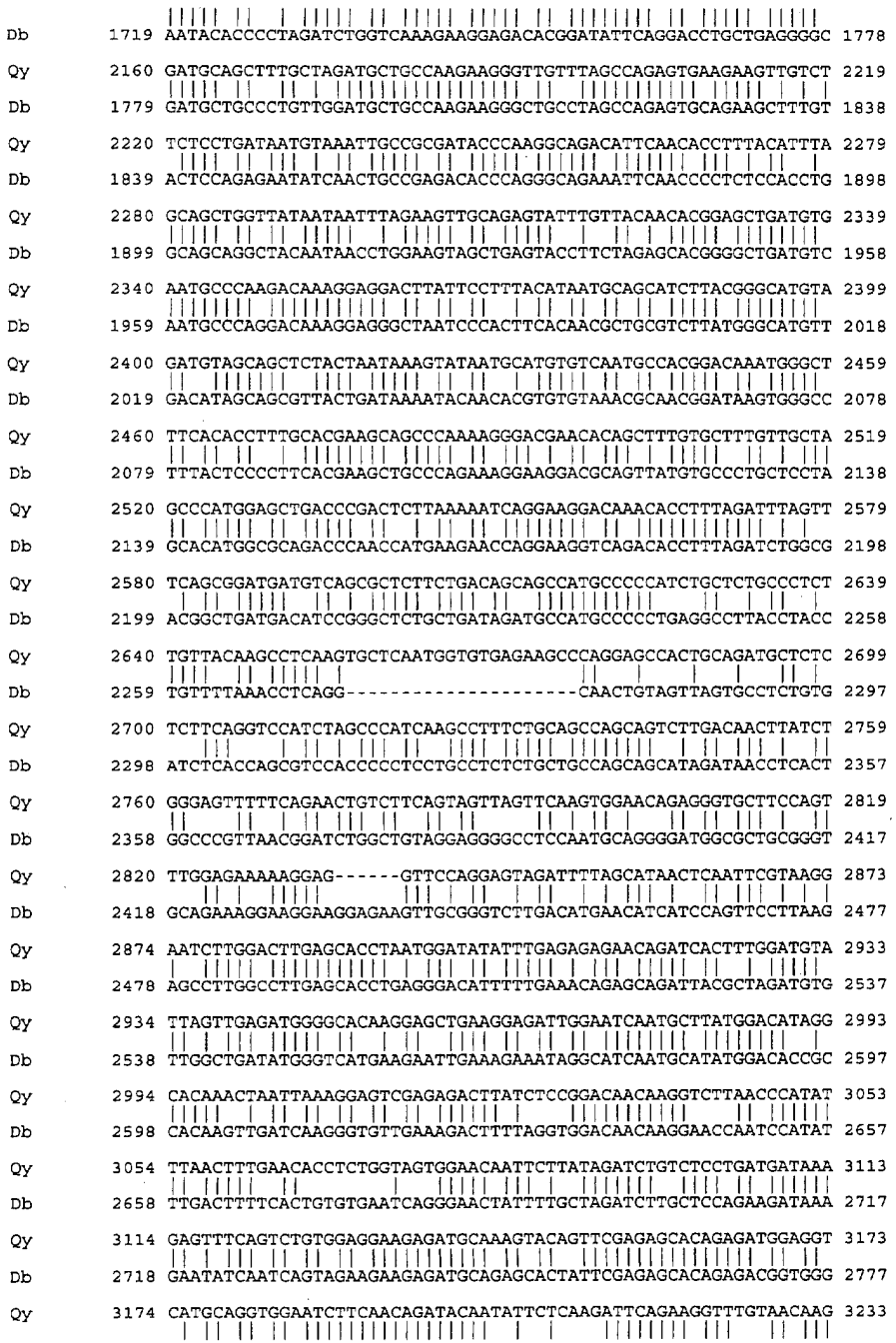
Qy       1860 CTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAAT 1919
              |||||  ||||||||||||||||  |||||  |||||  |||||  |||  |  ||  ||  |||||||
Db       1479 CTACATCATGGAGCTGATGTCCATGCCAAAGACAAAGGTGGCTTGGTTCCCCTTCACAAT 1538

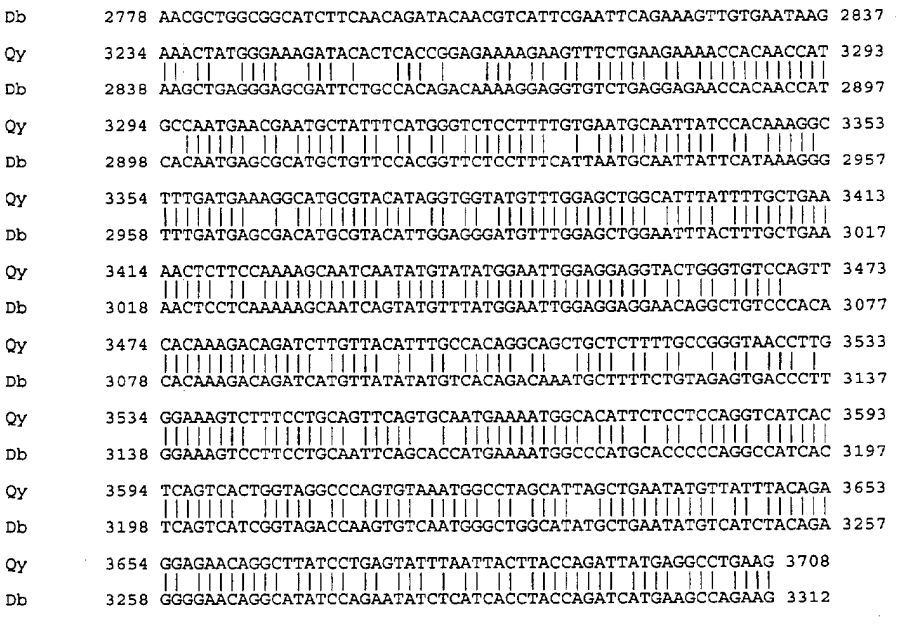
Qy       1920 GCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTT 1979
              ||  ||  ||  ||||||||  ||  ||  ||  ||  |  ||  ||  ||  ||  |||
Db       1539 GCCTGCTCCTATGGACACTATGAGGTCGCCGAGCTCTTGGTGAGGCACGGGGCTTCAGTC 1598

Qy       1980 AATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATAT 2039
              |||||  ||  ||  |  ||||||||  ||  |||  |  ||||||||  |||||  ||  ||||||||
Db       1599 AATGTGGCCGACCTGTGGAAATTCACCCCTCTCCATGAAGCGGCAGCCAAGGGAAAATAC 1658

Qy       2040 GAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGA 2099
              |||||  ||||||||  ||  ||  |  |||||  |||||  ||  ||  ||  |||||  ||||||
Db       1659 GAAATCTGCAAACTCCTTCTAAAACATGGAGCAGATCCAACTAAGAAGAACAGAGATGGA 1718

Qy       2100 AATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGA 2159
```





RESULT 3
AK080913
LOCUS AK080913 3194 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length enriched library, clone:B430206N15 product:TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE homolog [Homo sapiens], full insert sequence.
ACCESSION AK080913
VERSION AK080913.1 GI:26099551
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

```
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3194)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .3194
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:B430206N15"
                     /db_xref="MGI:2412078"
                     /db_xref="taxon:10090"
                     /clone="B430206N15"
                     /sex="male"
                     /tissue_type="adipose"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="4 days neonate"
     misc_feature    27. .>3193
                     /note="TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE
                     POLYMERASE homolog [Homo sapiens] (SPTR|O95271, evidence:
                     FASTY, 95.9%ID, 80%length, match=3167)
                     putative"
ORIGIN

  Query Match             37.6%;  Score 1436.2;  DB 11;  Length 3194;
  Best Local Similarity   71.8%;  Pred. No. 5.8e-287;
  Matches 1934;  Conservative    0;  Mismatches  733;  Indels   28;  Gaps    3;

Qy        263 CGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAGCTGTTCGAGGCGTGCCG 322
Db        519 CTCGGCCCCGGGGGTCCCAGCCGTGAGTGGGGCCCTGCGGGAGCTGCTCGAGGCCTGTCG 578

Qy        323 CAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAGAAGGTGAACAGCCGCGA 382
Db        579 GAATGGGGACGTGTCCCGGGTGAAGAGGTTGGTGGACGCCGCGAACGTGAATGCCAAGGA 638

Qy        383 CACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGTTTTGGGCGGAAAGACGT 442
Db        639 CATGGCCGGCCGAAAGTCTTCTCCCCTGCACTTCGCAGCAGGTTTTGGAAGGAAGGATGT 698

Qy        443 AGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAG-CACGTGATGATGGGGGCCTTA 501
Db        699 TGTACAACACTTGCTGCAGATGGGTGCTAATGTCCACGCCCCGTGATGACGGGGGTCTCA 758

Qy        502 TTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGAC 561
Db        759 TCCCACTTCATAATGCCTGCTCTTTCGGCCATGCTGAGGTTGTGAGTCTGTTACTGTGTC 818

Qy        562 ATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAA 621
Db        819 AAGGAGCAGATCCGAATGCCAGAGATAACTGGAACTATACACCACTGCATGAAGCTGCCA 878

Qy        622 TTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCC 681
Db        879 TTAAAGGGAAAATTGATGTCTGTATTGTGCTGCTGCAGCACGGAGCCGATCCAAACATCC 938

Qy        682 GAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTA 741
Db        939 GGAACACCGATGGGAAATCAGCTCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTCA 998

Qy        742 CTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAA 801
Db        999 CAGGTGAATACAAAAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAAC 1058

Qy        802 TGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAA 861
Db        1059 TAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGACGAAAGTCTA 1118

Qy        862 CTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAAC 921
Db        1119 CTCCGTTACACCTGGCAGCAGGCTACAACAGAGTTCGGATAGTTCAGCTGCTTCTACAGC 1178

Qy        922 ATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTT 981
Db        1179 ATGGTGCTGATGTCCACGCGAAAGACAAAGGTGGACTTGTACCTCTTCATAATGCATGCT 1238

Qy        982 CTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAA 1041
Db        1239 CATACGGACACTATGAGGTCACAGAACTGCTGCTAAAGCATGGAGCTTGTGTCAACGCTA 1298

Qy        1042 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 1101
Db        1299 TGGATCTGTGGCAGTTCACACCGCTGCATGAAGCTGCTTCAAAGAACCGGGTAGAAGTCT 1358

Qy        1102 GTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 1161
Db        1359 GTTCTCTGTTACTTAGTCACGGTGCCGATCCAACCTTGGTCAATTGCCATGGCAAAAGTG 1418

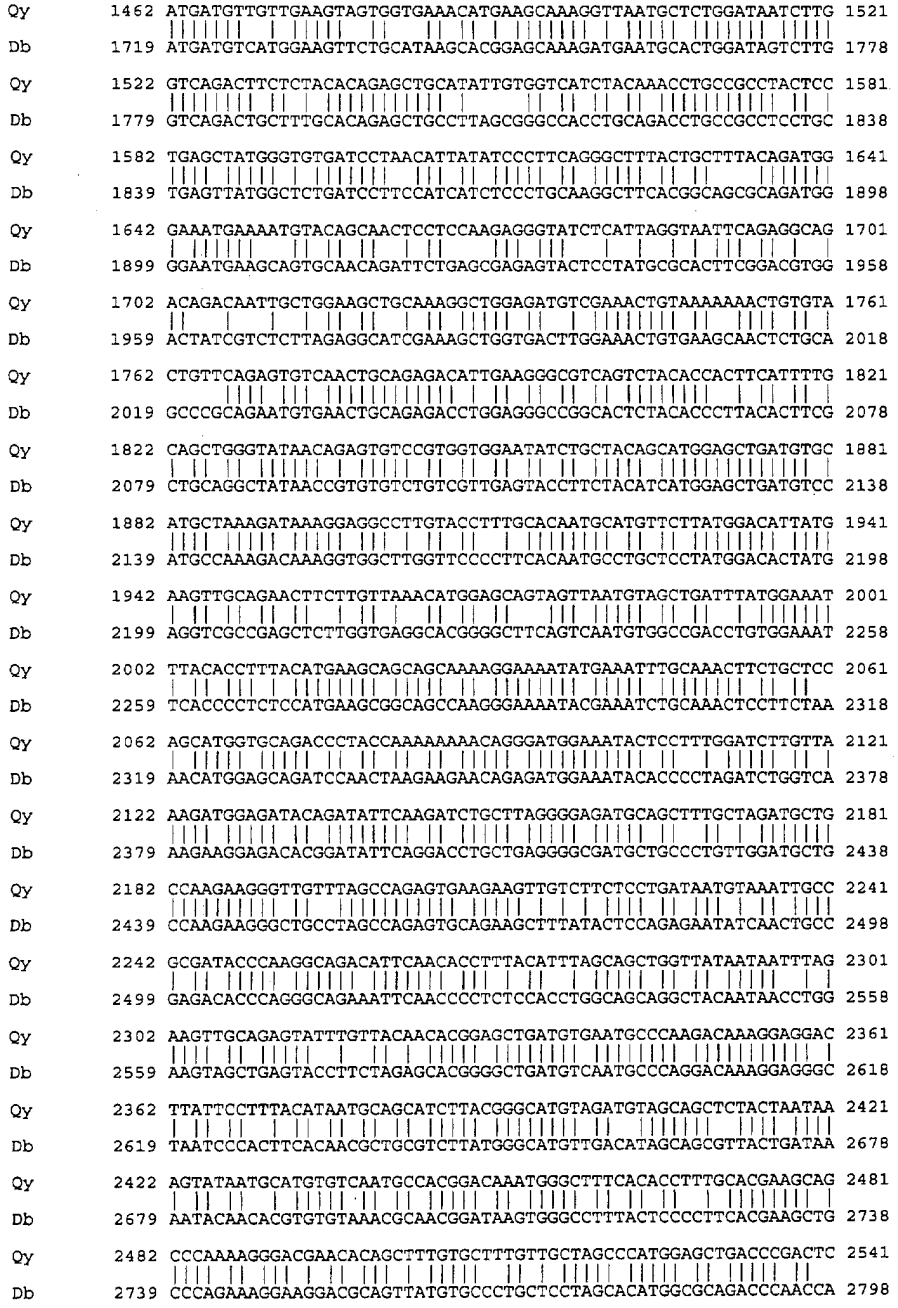
Qy        1162 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 1221
Db        1419 CTGTCGACATGGCTCCAACTCCTGAGCTCCGGGAAAGACTGACATATGAATTTAAAGGCC 1478

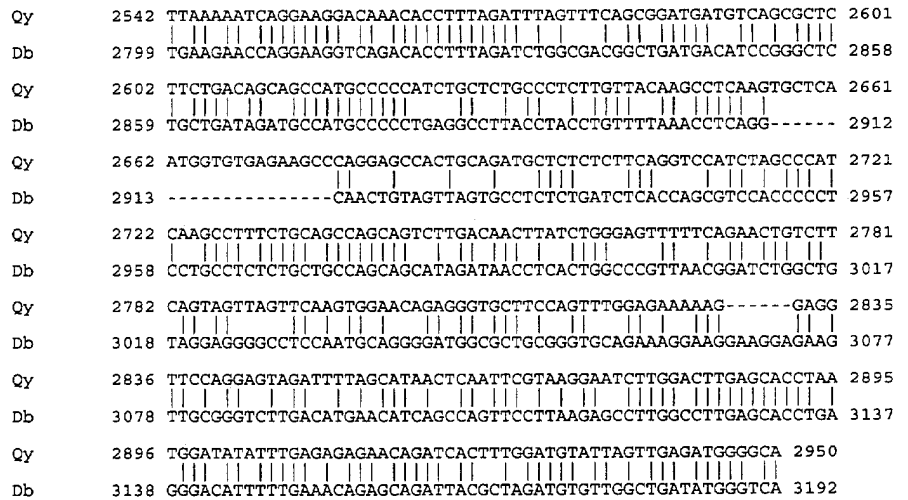
Qy        1222 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 1281
Db        1479 ATTCTTTACTCCAAGCCGCCAGAGAAGCAGACCTGGCCAAAGTTAAAAAAACGCTTGCTT 1538

Qy        1282 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 1341
Db        1539 TGGAAATCATTAACTTCAAGCAGCCACAGTCTCATGAAACTGCATTGCACTGTGCTGTGG 1598

Qy        1342 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 1401
Db        1599 CCTCCTTGCATCCCAAACGAAAGCAGGTGGCAGAGCTGTTACTGAGGAAGGGAGCCAACG 1658

Qy        1402 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 1461
Db        1659 TCAACGAAAAGAACAAGGATTTCATGACCCCTCTGCATGTTGCAGCTGAAAGAGCTCATA 1718
```





```
RESULT 4
AU140145
LOCUS       AU140145                838 bp    mRNA    linear   EST 05-AUG-2002
DEFINITION  AU140145 PLACE2 Homo sapiens cDNA clone PLACE2000021 5', mRNA
            sequence.
ACCESSION   AU140145
VERSION     AU140145.1  GI:11001666
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
            Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
            Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
            Masuho,Y., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1..838
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE2000021"
                     /tissue_type="placenta"
                     /clone_lib="PLACE2"
                     /note="Vector: pME18SFL3"
ORIGIN

  Query Match           20.6%;  Score 784.2;  DB 9;  Length 838;
  Best Local Similarity   98.0%;  Pred. No. 7.9e-152;
  Matches  823;  Conservative    0;  Mismatches   13;  Indels    4;  Gaps    3;
```

```
Qy       2734 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 2793
Db          1 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 60

Qy       2794 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 2853
Db         61 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 120

Qy       2854 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 2913
Db        121 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 180

Qy       2914 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 2973
Db        181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 240

Qy       2974 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 3033
Db        241 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 300

Qy       3034 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 3093
Db        301 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 360

Qy       3094 ATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 3153
Db        361 ATCTGTCTCCTGATGATAAAGAGTTTCCGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 420

Qy       3154 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 3213
Db        421 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 480

Qy       3214 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3273
Db        481 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 540

Qy       3274 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3333
Db        541 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 600

Qy       3334 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3393
Db        601 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 660

Qy       3394 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3453
Db        661 CTGGCATTTATTTTGCTGAAAACTCTT-CAAAAGCAATCAATATGTATATGGAATTGGAN 719

Qy       3454 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3513
Db        720 GANGT-CTGGGTGTCCAGTTCACAAAGACCGATCTTGGTACATTTGCCACANGCAGCTGC 778

Qy       3514 TCTTTTG--CCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 3571
Db        779 TCTTTTGGCCGGGTAACCCTTGGGAAAGTCTTTTCTGCAGTTCANTGCAATGNAAATGGC 838
```

RESULT 5
CB183123
LOCUS CB183123 955 bp mRNA linear EST 31-JAN-2003
DEFINITION AGENCOURT_11385409 NIH_MGC_164 Mus musculus cDNA clone IMAGE:30244272 5', mRNA sequence.
ACCESSION CB183123
VERSION CB183123.1 GI:28181571
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 955)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0316 row: o column: 01
High quality sequence stop: 779.
FEATURES Location/Qualifiers
source 1. .955
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30244272"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_164"
/note="Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Non-normalized full-length enriched library from pooled mouse embryonic limb, maxilla and mandible, day 10.5 and 11.5 (size selected for the 0.5-1 kb fragments) Cloned directionally, priming method: Oligo-dT. cDNA enrichment: >1k bp, Average insert size 1.8k bp. Priming sequence: 5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T) 3'. Tissue contributed by, David Rowe. Library constructed by ResGen, Invitrogen Corp."
ORIGIN

Query Match 18.0%; Score 687.8; DB 14; Length 955;
Best Local Similarity 88.5%; Pred. No. 8.8e-132;
Matches 768; Conservative 0; Mismatches 98; Indels 2; Gaps 2;

```
Qy       2835 GTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTA 2894
Db          1 GATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACTTG 60

Qy       2895 ATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAG 2954
Db         61 ATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAAG 120

Qy       2955 GAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTC 3014
Db        121 GAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACAAGCTGATTAAAGGAGTT 180

Qy       3015 GAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGT 3074
Db        181 GAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAACTCTAAACAACTCTGGT 240

Qy       3075 AGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAA 3134
Db        241 AGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAA 300

Qy       3135 GAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAAC 3194
Db        301 GAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAAC 360

Qy       3195 AGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACT 3254
Db        361 AGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTGGGAAAGATATACA 420

Qy       3255 CACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTT 3314
Db        421 CACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCAAACGAAAGGATGTTATTT 480

Qy       3315 CATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTAC 3374
Db        481 CATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTGATGAAAGGCATGCATAC 540

Qy       3375 ATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAA 3434
```

```
Db        541 ATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTCCAAAAGCAATCAA 600

Qy       3435 TATGTATAT-GGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTA 3493
              ||||| ||| |||||||||| || || || |||||| ||||||||||||||| |||||
Db        601 TATGTGTATNGGAATTGGAGGTGGCACCGGATGTCCAATTCACAAAGACAGATCGTGTTA 660

Qy       3494 CATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTT 3553
              |||||| |||||||||||||| ||||| || |||||||||||||||||||||| |||||||
Db        661 CATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAAAGTCTTTCTTGCAGTT 720

Qy       3554 CAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAG 3613
              ||| |||||||||||||||||||||||| || |||||||| |||||||  |||||||
Db        721 CAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGGTCACTGGCCGGCCCAG 780

Qy       3614 TGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGA 3673
              ||||||| |||||||||||||||||||||||||||||||||||||||||||    |||||
Db        781 TGTAAAT-GCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTATCCCTGA 839

Qy       3674 GTATTTAATTACTTACCAGATTATGAGG 3701
                  | | |||  |  || | | || |
Db        840 TATCTATACTACCAATGAGCTGANGATG 867
```

```
RESULT 6
CB128566
LOCUS       CB128566                675 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0178055 L11SNU354 Homo sapiens cDNA clone L11SNU354-7-A10 5',
            mRNA sequence.
ACCESSION   CB128566
VERSION     CB128566.1  GI:28091690
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 7  row: A  column: 10
            High quality sequence stop: 675.
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L11SNU354-7-A10"
                     /sex="M"
                     /tissue_type="Liver"
                     /cell_type="Polygonal"
                     /cell_line="SNU-354"
                     /lab_host="Top10F'"
                     /clone_lib="L11SNU354"
                     /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN
```

```
  Query Match             17.6%;  Score 672.4;  DB 14;  Length 675;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-128;
  Matches  673;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2436 GTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGA 2495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGA 60

Qy       2496 ACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAA 2555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAA 120

Qy       2556 GGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCC 2615
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        121 GGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCC 180

Qy       2616 ATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGC 2675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGC 240

Qy       2676 CCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCA 2735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCA 300

Qy       2736 GCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCA 2795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCA 360

Qy       2796 AGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGC 2855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGC 420

Qy       2856 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 2915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 480

Qy       2916 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 2975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 540

Qy       2976 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 3035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 600

Qy       3036 CAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 3095
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        601 CAAGGTCTTAACCCATATTTAACTNTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 660

Qy       3096 CTGTCTCCTGATGAT 3110
              ||||||||||||||||
Db        661 CTGTCTCCTGATGAT 675
```

```
RESULT 7
BM457025
LOCUS       BM457025                1128 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6406851 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583257
            5', mRNA sequence.
ACCESSION   BM457025
VERSION     BM457025.1  GI:18506065
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1128)
```

```
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12346  row: a  column: 18
            High quality sequence stop: 507.
FEATURES             Location/Qualifiers
     source          1. .1128
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5583257"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             17.6%;  Score 672.2;  DB 12;  Length 1128;
  Best Local Similarity   97.7%;  Pred. No. 1.7e-128;
  Matches  724;  Conservative    0;  Mismatches   13;  Indels    4;  Gaps    4;

Qy       1976 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 2035
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 60

Qy       2036 ATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGA 2095
             |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 ATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGA 120

Qy       2096 TGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 2155
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 180

Qy       2156 GGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTT 2215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTT 240

Qy       2216 GTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACA 2275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACA 300

Qy       2276 TTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGA 2335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGA 360

Qy       2336 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCA 2395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCA 420

Qy       2396 TGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATG 2455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATG 480

Qy       2456 GGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTT 2515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTT 540

Qy       2516 GCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA-CCTTTAGATT 2574
             |||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db        541 GCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCCTTTACATT 600

Qy       2575 TAGTTTCAGCGGATGATGTCAGCG-CTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTG 2633
             |||||||||||||||||||||||| |||||||||| |||||||| ||||||||||||||||
Db         601 TAGTTTCAGCGGATGATGTCAGCGCCTCTTCTGACACCAGCCATGGCCCCATCTGCTCTG 660

Qy       2634 CCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTG-TGAGAAGCCCAGGAGCCACTGCAGA 2692
             |||||||| ||| |||||||| |||| |||||| |||||||||  ||||||||||||||||
Db         661 CCCTCTTGGTACCAGCCTCAAGCGCTCCATGGTGCTGAGAAGCCACGGAGCCACTGCAGA 720

Qy       2693 TG-CTCTCTCTTCAGGTCCAT 2712
             || || |||||||||||| | |
Db         721 TGCCTTTCTCTTCAGGTTCCT 741


RESULT 8
CB170279/c
LOCUS       CB170279                754 bp    mRNA    linear   EST 30-JAN-2003
DEFINITION  CEY603161575.R1 CSEQFXN42 testes Bos taurus cDNA, mRNA sequence.
ACCESSION   CB170279
VERSION     CB170279.1  GI:28156407
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   Adelson,D.L. and Gill,C.A.
  TITLE     Bovine ESTs (Adelson and Gill)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: David L. Adelson
            Animal Breeding and Genetics
            Texas A&M University
            Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
            USA
            Tel: 9798452616
            Fax: 9798456970
            Email: david.adelson@tamu.edu.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="testes"
                     /clone_lib="CSEQFXN42 testes"
                     /note="Organ: testes; Vector: pBluescript SK+; Site_1:
                     NotI; Site_2: EcoRI; sequence 5' of the insert
                     (5'-NNN. . .NNNinsert)
                     GCGAATTGGAGCTCCACCGCGGTGGCGGCCGCGGCTCGAG. Sequence 3' of
                     the inserts (AAGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG.
                     normalized Rd 2 library, sequenced 3' with M13R primer."
ORIGIN

  Query Match             17.4%;  Score 665.2;  DB 14;  Length 754;
  Best Local Similarity   93.5%;  Pred. No. 3.9e-127;
  Matches  705;  Conservative    0;  Mismatches   48;  Indels    1;  Gaps    1;

Qy        610 ATGAAGCTGCAATTAAAGGAAAGATTGATGTTTG-CATTGTGCTGTTACAGCATGGAGCT 668
              |||||||||||||||||||||||||||||||||| || |||||||||| ||||||||||||||
Db        754 ATGAAGCTGCAATTAAAGGAAAGATTGATGTCTGCCATTGTGCTCTTACAGCATGGAGCT 695

Qy        669 GAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCC 728
              |||||||||| |||||||||||||| |||||||||||||||||||||| || ||||||
Db        694 GAGCCAACCATTCGAAATACAGATGGACGGACAGCATTGGATTTAGCAGACCCGTCTGCC 635

Qy        729 AAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGC 788
              |||||||||| |||||||||||||||| |||||||||| |||||||||||||||||||||
Db        634 AAAGCAGTGCTGACTGGTGAATATAAGAAGGATGAACTCTTGGAAAGTGCCAGGAGTGGC 575

Qy        789 AATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGAT 848
              ||||| |||||||||||||||| || |||||||||||||| ||||| || ||||||||||
Db        574 AATGAGGAAAAAATGATGGCTCTGCTTACACCATTAAATGTAAACTGTCATGCAAGTGAT 515
```

```
Qy       849 GGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAG 908
Db       514 GGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGTTACAACAGAGTAAAGATTGTACAG 455

Qy       909 CTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTA 968
Db       454 CTGTTGTTGCAACATGGAGCTGATGTTCATGCTAAAGATAAAGGTGATCTGGTACCCCTG 395

Qy       969 CACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCC 1028
Db       394 CACAACGCCTGTTCCTATGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCATGGTGCC 335

Qy      1029 TGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAAC 1088
Db       334 TGCGTAAATGCAATGGACTTGTGGCAGTTCACTCCTCTCCATGAGGCTGCTTCCAAGAAC 275

Qy      1089 AGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGT 1148
Db       274 AGGGTTGAAGTGTGTTCTCTTCTCCTGAGTTATGGCGCAGATCCAACACTGCTGAACTGT 215

Qy      1149 CACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATAT 1208
Db       214 CACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGGTTAGCATAT 155

Qy      1209 GAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAA 1268
Db       154 GAATTTAAAGGCCACTCCTTGCTTCAGGCTGCACGGGAAGCTGATGTTACACGAATCAAA 95

Qy      1269 AAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTG 1328
Db        94 AAACATCTCTCTCTGGAAATGGTGAACTTTAAGCATCCTCAAACACATGAAACAGCATTG 35

Qy      1329 CATTGTGCTGCTGCATCTCCATATCCCAAAAGAA 1362
Db         34 CATTGTGCTGCTGCATCTCCATATCCCAAAAAAA 1
```

RESULT 9
CF407574/c
LOCUS CF407574 923 bp mRNA linear EST 02-SEP-2003
DEFINITION CH3#049_C04T3 Canine heart normalized cDNA Library in pBluescript Canis familiaris cDNA clone CH3#049_C04 3', mRNA sequence.
ACCESSION CF407574
VERSION CF407574.1 GI:34408600
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 923)
AUTHORS Yi,Y., Desai,R., Olarte,M., Henthorn,P. and George A.L.
TITLE Expressed sequence tags from Canine heart
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: CH3#049_C04T7
Contact: George AL
Division of Genetic Medicine
Vanderbilt University
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
Tel: 615 936 2660
Fax: 615 936 2661
Email: al.george@vanderbilt.edu
Insert Length: 2028 Std Error: 0.00
Seq primer: T3: ATTAACCCTCACTAAAGGGA
High quality sequence start: 65
High quality sequence stop: 838.
FEATURES Location/Qualifiers
source 1. .923
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CH3#049_C04"
/tissue_type="heart"
/cell_type="heart"
/dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
/clone_lib="Canine heart normalized cDNA Library in pBluescript"
/note="Organ: heart; Vector: pBluescript; Site_1: 5' of vector NotI; Site_2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"
ORIGIN

Query Match 17.1%; Score 652; DB 14; Length 923;
Best Local Similarity 94.3%; Pred. No. 2.4e-124;
Matches 676; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy      2603 TCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAA 2662
Db       893 TCTGACAGCAGCCATGCCCCCTTNTGCTCTGCCTTCATGTTATAAACCTCAAGTGCTCAA 834

Qy      2663 TGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATC 2722
Db       833 TGGCGTGAGAAGCCCAGGAGCTACCGGCGATGCTCTGTCTTCAGGTCCATCCAGCCCATC 774

Qy      2723 AAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTC 2782
Db       773 AAGTCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCTGAACTGTCTTC 714

Qy      2783 AGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGG 2842
Db       713 AGTAGTTAGTTCAAGTGGAACAGAGGGTGCCTCTAGTTTGGAGAAAAAGGAGGTTCCAGC 654

Qy      2843 AGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATAT 2902
Db       653 AGTAGATTTTAGCATAACTCAGTTTGTAAGGAATCTTGGGCTTGAGCACCTGATGGATAT 594

Qy      2903 ATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAA 2962
Db       593 ATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTGGAGATGGGGCACAAGGAGCTGAA 534

Qy      2963 GGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACT 3022
Db       533 GGAGATTGGAATCAACGCTTATGGACATAGACACAAACTAATTAAAGGAGTTGAGAGACT 474

Qy      3023 TATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAAC 3082
Db       473 TATATCTGGACAACAAGGTCTTAACCCATACTTAACTCTGAACACCTCTGGAAGTGGAAC 414

Qy      3083 AATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCA 3142
Db       413 AATTCTCATAGATCTATCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCA 354

Qy      3143 AAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAA 3202
Db       353 AAGTACAGTCCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAA 294

Qy      3203 TATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAG 3262
Db       293 TATTCTCAAGATTCAGAAAGTTTGTAACAAGAAATTATGGGAAAGATATACACACAGGAG 234

Qy      3263 AAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3319
Db       233 AAAAGAAGTTTCTGAAGAAAACCACAACCATGCAAATGAAAGAATGCTATTTCATGG 177
```

RESULT 10
R64714
LOCUS R64714 659 bp mRNA linear EST 12-JUN-1996
DEFINITION EST54a09 WATM1 Homo sapiens cDNA clone 54a09 similar to BLASTX homology with sp|Q01485|ANKC_HUMAN ANKYRIN, BRAIN VARIANT 2, mRNA sequence.
ACCESSION R64714
VERSION R64714.1 GI:837352

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Bouillaud,F.
  TITLE     Study of expressed sequences tags in adipose tissue 1995
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Frederic Bouillaud
            Centre de Recherche sur l'Endocrinologie moleculaire et le
            Developpement
            CNRS
            9, Rue Jules Hetzel, Meudon Bellevue, 92190 France
            Tel: 33 1 45 07 52 87
            Fax: 33 1 45 07 58 90
            Email: bouillau@infobiogen.fr
            automatic cycle sequencing of PCR amplified insert, open reading
            frame 3. .659.
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /strain="Caucasian"
                     /db_xref="taxon:9606"
                     /clone="54a09"
                     /clone_lib="WATM1"
                     /note="Vector: lambda gt11; Site_1: EcoRI; White adipose
                     tissue, sub cutaneous, adult, female. Purification of
                     polyA mRNA, first strand priming with random
                     oligonucleotides. Ligation of adaptors EcoRI/NotI, cloning
                     in EcoRI site of lambda gt11."
ORIGIN

  Query Match            17.0%;  Score 649.2;  DB 14;  Length 659;
  Best Local Similarity  98.6%;  Pred. No. 7.7e-124;
  Matches  649;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1435 TGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAG 1494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAG 60

Qy       1495 CAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTG 1554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db         61 CAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATHTTGTG 120

Qy       1555 GTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCC 1614
             |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GYCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCC 180

Qy       1615 TTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGG 1674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGG 240

Qy       1675 GTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAG 1734
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        241 GTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTSCAAAGGCTGGAG 300

Qy       1735 ATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAG 1794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAG 360

Qy       1795 GGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAT 1854
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAT 420

Qy       1855 ATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGC 1914
             ||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCTGCYACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGC 480

Qy       1915 ACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAG 1974
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAG 540

Qy       1975 TAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAA 2034
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |
Db         541 TAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCCAAAGGGA 600

Qy       2035 AATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAG 2092
             ||:|||||||||| ||||||||||||||||||||||||||||| ||||||||||||
Db         601 AAWATGAAATTTGCCAACTTCTGCTCCAGCATGGTGCAGACCCCACCAAAAAAAACAG 658


RESULT 11
BG391376
LOCUS       BG391376                691 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602417447F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4537138 5',
            mRNA sequence.
ACCESSION   BG391376
VERSION     BG391376.1  GI:13284824
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10461  row: m  column: 11
            High quality sequence stop: 684.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4537138"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            16.9%;  Score 643.8;  DB 12;  Length 691;
  Best Local Similarity  99.3%;  Pred. No. 1e-122;
  Matches  678;  Conservative    0;  Mismatches    2;  Indels    3;  Gaps    3;

Qy       1725 AAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          10 AAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 69

Qy       1785 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          70 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 129

Qy       1845 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1904
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 189

Qy       1905 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1964
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 249
```

```
Qy       1965 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2024
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 309

Qy       2025 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCA-TGGTGCAGACCCTACCAA 2083
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||
Db        310 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCACTGGTGCAGACCCTACAAA 369

Qy       2084 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 2143
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 429

Qy       2144 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 2203
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 489

Qy       2204 AGTGAAGAAGTTGTCTTCTCCTGATA-ATGTAAATTGCCGCGATACCCAAGGCAGACATT 2262
             |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        490 AGTGAAGAAGTTGTCTTCTCCTGATACATGTAAATTGCCGCGATACCCAAGGCAGACATT 549

Qy       2263 CAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTAC 2322
             |||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db        550 CAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA-TAGTTAC 608

Qy       2323 AACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 AACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 668

Qy       2383 CATCTTACGGGCATGTAGATGTA 2405
             |||||||||||||||||||||||
Db        669 CATCTTACGGGCATGTAGATGTA 691
```

```
RESULT 12
AL601027
LOCUS       AL601027                759 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp313G2239_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
            DKFZp313G2239 5', mRNA sequence.
ACCESSION   AL601027
VERSION     AL601027.1  GI:15164533
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Duesterhoeft, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp313G2239) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313G2239"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="313 (synonym: hlcc2)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
```

```
ORIGIN

  Query Match           16.8%;  Score 642.8;  DB 9;  Length 759;
  Best Local Similarity  97.2%;  Pred. No. 1.8e-122;
  Matches  669;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    2;

Qy       2393 GCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAA 2452
             ||||||| ||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db         72 GCATGTANATGTAGCAGCTCTACTAATAAAGNNTAATGCATGTGTCAATGCCACGGGCAA 131

Qy       2453 ATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTT 2512
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        132 ATGGGCTTTCACACCTTTGCACGAAGCAGCCCAGAAGGGACGAACACAGCTTTGTGCTTT 191

Qy       2513 GTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGA 2572
              ||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTTTTTTGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGA 251

Qy       2573 TTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCT 2632
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCT 311

Qy       2633 GCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGA 2692
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 GCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGA 371

Qy       2693 TGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAA 2752
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        372 TGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGNCTTGACAA 431

Qy       2753 CTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGC 2812
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGC 491

Qy       2813 TTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAG 2872
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        492 TTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGNAAG 551

Qy       2873 GAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 2932
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 611

Qy       2933 ATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAG 2992
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 ANNAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAG 671

Qy       2993 GCACAAACTAATTAAA-GGAGTCGAGAGACTTATCTCCGGACAACAA-GGTCTTAACCCA 3050
             |||||||||||||||| |||||||||||||| |||||||||||||| || |||||||||
Db        672 GCACAAACTAATTAAAGGGAGTCGAGAGACTNATCTCCGGACAACAAGGGGCTTAACCCA 731

Qy       3051 TATTTAACTTTGAACACCTCTGGTAGTG 3078
             |||||||||||||||||||||||| || |
Db        732 TATTTAACTTTGAACACCTCTGGNAGAG 759
```

```
RESULT 13
BQ885764
LOCUS       BQ885764                950 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8746924 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6392888
            5', mRNA sequence.
ACCESSION   BQ885764
VERSION     BQ885764.1  GI:22277782
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 950)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM13883 row: h column: 09
High quality sequence stop: 641."

```
FEATURES             Location/Qualifiers
     source          1. .950
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6392888"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_130"
                     /note="Organ: otocysts; Vector: pCMV-SPORT6.1; Site_1:
                     EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 1.95 kb. Constructed by
                     ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             16.6%;  Score 634.4;  DB 13;  Length 950;
  Best Local Similarity   86.4%;  Pred. No. 1.1e-120;
  Matches  759;  Conservative    0;  Mismatches  111;  Indels    8;  Gaps    5;

Qy       2500 AGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGAC 2559
             ||||||||||||| ||| | |||||||||||||| || |||||| ||||||| ||||||
Db          1 AGCTTTGTGCTTTATTGTTGGCCCATGGAGCTGATCCTACTCTT-AAAATCAAGAAGGAC 59

Qy       2560 AAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGC 2619
             ||||||||||||||||||||||||||| |||||||||||| || ||  ||||||||||||||
Db         60 AAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGC 119

Qy       2620 CCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAG 2679
             ||||  ||||||||||   | || ||||| |||||||||||||| ||| |||||  |||| |
Db        120 CCCCCTCTGCTCTGCCTACGTGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCG 179

Qy       2680 GAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCA 2739
             |||||||||||||||||||| |||||||||| || |||||||| |||||  ||||||||||
Db        180 GAGCCACTGCAGATGCTCTGTCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCA 239

Qy       2740 GCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTG 2799
             ||||  || |||||||||||||| || || || ||||||||   |||| ||||||||||||
Db        240 GCAGCCTCGACAACTTATCTGGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTG 299

Qy       2800 GAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAA 2859
               | |||| |||||| |  ||||| | | ||| |||| | |||||| | |||||||| ||||
Db        300 CAGCAGAAGGTGCTACTGGTTTGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAA 359

Qy       2860 CTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGA 2919
             ||||  ||| |||||||  |||||||||||||||| | ||||||||||||||||||||||||
Db        360 CTCAGTTCATAAGGAACCTTGGACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGA 419

Qy       2920 TCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATG 2979
             | || || |||||  |||||||  |||| |||||||| |||||  ||||||||||||||||
Db        420 TTACCTTAGATGTCCTAGTTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATG 479

Qy       2980 CTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAG 3039
             | |||||||||  | |||||  || |||||||||| || ||  ||  ||||| ||||||||||
Db        480 CCTATGGACATCGACACAAGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAACAAG 539

Qy       3040 GTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGT 3099
             |||||||  |||||||||||| | ||||  |||||||||||||||||||  ||||||||||
Db        540 GTCTTAATCCATATTTAACTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGT 599

Qy       3100 CTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGC 3159
             |||||||||||||||  |||||||||||||| ||||||||||  |||||  ||  ||  ||||
Db        600 CTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGC 659

Qy       3160 ACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGA 3219
             |||||||||| |||||  |||||||||||||||  |||||||||||||||  ||||||||||||
Db        660 ACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGA 719

Qy       3220 AGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG-AGAAAAGAAGTTTCTGAA 3278
             |||||||||||||||||  | |||||||||||  || |||||  ||||||||||||||||||
Db        720 AGGTTTGTAACAAGAAATTGTGGGAAAGATATACACACCGGNAGAAAAGAAGTTTCTGAA 779

Qy       3279 GAAAACCACAACCATGC-CAATGAACGAATGCTATTTCATGGGTCT---CCTTTTGTGAA 3334
             |||||||||||||||||  || ||| | ||| |||||    ||  |   | |||||||||
Db        780 GAAAACCACAACCATGCAAAACGAAAGGATGTTATTTTCATGGGGTCTCCTTTTTGTGAA 839

Qy       3335 TGCAATTATCCACAAAGG--CTTTGATGAAAGGCATGC 3370
             ||| |||||||| |||||   ||||||||||||||||||
Db        840 TGCGATTATCCATAAAGGGCTTTTGATGAAAGGCATGC 877
```

```
RESULT 14
CF539235
LOCUS       CF539235                743 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-GH0-chd-h-19-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
            IMAGE:30535386 5', mRNA sequence.
ACCESSION   CF539235
VERSION     CF539235.1  GI:34591604
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30535386"
                     /tissue_type="Whole brain"
                     /dev_stage="1, 5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GH0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match             16.6%;  Score 632.6;  DB 14;  Length 743;
  Best Local Similarity   90.7%;  Pred. No. 2.3e-120;
```

```
  Matches  674;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy       1705 GACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTG 1764
             ||||  ||||||||||||||||||||||| || ||  | ||||||||||||||| |||||||
Db          1 GACAGCTGCTGGAAGCTGCAAAGGCTGGTGACGTTGTAACTGTAAAAAAACTCTGTACTG 60

Qy       1765 TTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAG 1824
             |||||||||||||||||||||||||||||| || ||||| || ||||| || ||||||| ||
Db         61 TTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCAGTCAACCCCACTCCACTTTGCAG 120

Qy       1825 CTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATG 1884
             ||||||| |||||||||||  ||||| |||||||||| || ||||||||||||||| ||||
Db         121 CTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAACATGGAGCTGATGTTCATG 180

Qy       1885 CTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAG 1944
             ||||||| ||||||| |||||||||| |||||||||||||||||||||| |||||||
Db         181 CTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGCATGTTCTTATGGACACTATGAAG 240

Qy       1945 TTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTA 2004
             |||||||||||||||| || |||||||||||||| ||||| ||||| || ||||| || |
Db         241 TTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCA 300

Qy       2005 CACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGC 2064
             |||||||||||||||| |||||||||||||||||||||||| |||||||||| |||||||
Db         301 CACCTTTACATGAAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGC 360

Qy       2065 ATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAG 2124
             |||||||||||||||| || |||||||| ||||| |||||||| ||||||||||||||||
Db         361 ATGGTGCAGACCCTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAG 420

Qy       2125 ATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCA 2184
             |||||||||||||||||||||||| ||||| || |||||||||||||| ||||||||||
Db         421 ATGGAGATACAGATATTCAAGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCA 480

Qy       2185 AGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCG 2244
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||| |
Db         481 AGAAGGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTG 540

Qy       2245 ATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAG 2304
             ||||||||||  | |||| |||||||||||||| ||||||||||||||||| || ||||
Db         541 ATACCCAAGGTCGGCATTCTACACCTTTACATTTGGCAGCTGGTTATAATAACTTGGAAG 600

Qy       2305 TTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTA 2364
             |||| |||||||| || ||||| |||||||||||||||||||| |||||||||||||| |
Db         601 TTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGACAAAGGAGGACTCA 660

Qy       2365 TTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGT 2424
             |||||||||||| |||||||| || || |||||||||||||||||||| | || |||||||
Db         661 TTCCTTTACATAACGCAGCATCGTATGGACATGTAGATGTAGCAGCTTTGCTGATAAAGT 720

Qy       2425 ATAATGCATGTGTCAATGCCACG 2447
             | |  || ||||||||||||||||
Db         721 ACATGGCGTGTGTCAATGCCACG 743


RESULT 15
BX474686
LOCUS       BX474686                638 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686J06171_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686J06171 5', mRNA sequence.
ACCESSION   BX474686
VERSION     BX474686.1  GI:31668891
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686J06171) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686J06171"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match            16.3%;  Score 622.8;  DB 13;  Length 638;
  Best Local Similarity  99.5%;  Pred. No. 2.3e-118;
  Matches  635;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1805 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1864
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 60

Qy       1865 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1924
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCAAG 120

Qy       1925 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1984
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 180

Qy       1985 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2044
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 240

Qy       2045 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATAC 2104
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         241 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATAC 300

Qy       2105 TCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGC 2164
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGC 360

Qy       2165 AGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCC 2224
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCC 420

Qy       2225 TGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC-AG 2283
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db         421 TGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCNAG 480

Qy       2284 CTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATG 2343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATG 540

Qy       2344 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATG 2403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATG 600

Qy       2404 TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAAT 2441
             |||||||||||||||||||||||||||||||||||||||
Db         601 TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAAT 638
```

```
RESULT 16
CF533694
LOCUS       CF533694                724 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-FY0-cgs-i-17-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
            IMAGE:30362032 5', mRNA sequence.
ACCESSION   CF533694
VERSION     CF533694.1  GI:34585662
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30362032"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_FY0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match           16.2%;  Score 616.4;  DB 14;  Length 724;
  Best Local Similarity   91.2%;  Pred. No. 5.2e-117;
  Matches  653;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy       1908 CCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACAT 1967
              ||||| |||||||||||||||||||||||| |||||||||||||||||||||||| || |||
Db          1 CCTTTACACAATGCATGTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCAT 60

Qy       1968 GGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCA 2027
              |||||||||||| ||||| ||||| || |||||| || |||||||||||||||| ||||||
Db         61 GGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACCTTTACATGAAGCTGCAGCA 120

Qy       2028 AAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAA 2087
              |||||||||||||||| |||||||||| |||||||||||||||||||||||||| || |||
Db        121 AAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGGTGCAGACCCTACAAAGAAA 180

Qy       2088 AACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGAT 2147
              ||||| ||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||

Db        181 AACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGAT 240

Qy       2148 CTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTG 2207
              || ||||| || |||||||||||||| ||||||||||||||||||||||||||||||||||
Db        241 CTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTG 300

Qy       2208 AAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACA 2267
              ||||| |||||||||||||||||||||||||||||| |||||||||||  | ||||| |||
Db        301 AAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACA 360

Qy       2268 CCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACAC 2327
              |||||||||| |||||||||||||||||| || |||||||| |||||||| || |||||
Db        361 CCTTTACATTTGGCAGCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACAT 420

Qy       2328 GGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCT 2387
              |||||||||||||||||||| |||||||||||||| |||||||||||||| ||||||||
Db        421 GGAGCTGATGTGAATGCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCG 480

Qy       2388 TACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACG 2447
              || || |||||||||||||||||| | || |||||||| ||||| ||||||||||||||||
Db        481 TATGGACATGTAGATGTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACG 540

Qy       2448 GACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGT 2507
              |||||||||||||||||| || || |||||||||||||||||||| ||||||||||||
Db        541 GACAAATGGGCTTTCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGT 600

Qy       2508 GCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCT 2567
              ||||| ||| | |||||||||||||| || ||||||| |||||| |||||||| ||||||
Db        601 GCTTTATTGTTGGCCCATGGAGCTGATCCTACTCTTANAAATCAAGAAGGACANACACCT 660

Qy       2568 TTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCC 2623
              |||||||||||||||| |||||||||| || ||  |||||||||||||||||||
Db        661 TTAGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCC 716


RESULT 17
CF727485
LOCUS       CF727485                747 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  UI-M-HB0-ckk-b-13-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
            IMAGE:30547140 5', mRNA sequence.
ACCESSION   CF727485
VERSION     CF727485.1  GI:37601653
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .747
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30547140"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HB0"
```

/note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 16.1%; Score 614.2; DB 14; Length 747;
Best Local Similarity 90.5%; Pred. No. 1.5e-116;
Matches 677; Conservative 0; Mismatches 69; Indels 2; Gaps 2;

```
Qy       1924 GTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATG 1983
Db          1 GTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATG 60

Qy       1984 TAGCTGATTTAT-GGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2042
Db         61 TCGCTGACTTGTNGGAAGTTCACACCTTTACATGAAGCTGCAGCAAAAGGAAAATATGAA 120

Qy       2043 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 2102
Db         121 ATATGCAAACTTCTACTCCAGCATGGTGCAGACCCTACAAAGAAAAACAGAGATGGTAAT 180

Qy       2103 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2162
Db         181 ACTCCTCTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTACTTAGAGGTGAT 240

Qy       2163 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2222
Db         241 GCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTCT 300

Qy       2223 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2282
Db         301 CCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACACCTTTACATTTGGCA 360

Qy       2283 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2342
Db         361 GCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAAT 420

Qy       2343 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2402
Db         421 GCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTAGAT 480

Qy       2403 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2462
Db         481 GTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTTC 540

Qy       2463 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2522
Db         541 ACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGCC 600

Qy       2523 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2582
Db         601 CATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACANACACCTTTAGATTTAGTTTCA 660

Qy       2583 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2642
Db         661 GCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGC 720

Qy       2643 TACAAGCCTCAAGTGCTCAATGGTGTGA 2670
Db         721 TACAA-ACTCAAGTGCTCAGTGGCGTGA 747
```

RESULT 18
BG875327
LOCUS BG875327 648 bp mRNA linear EST 30-MAY-2001
DEFINITION RC1-BT0623-120200-011-d04 BT0623 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG875327
VERSION BG875327.1 GI:14252306
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 648)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC1-BT0623-120200-011-d04&t3=2000-02-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 4
High quality sequence stop: 648.

FEATURES Location/Qualifiers
source 1. .648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT0623"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 15.9%; Score 605; DB 12; Length 648;
Best Local Similarity 98.7%; Pred. No. 1.2e-114;
Matches 620; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

```
Qy       1663 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1722
Db         23 TCCTACAAGAGGGTATCTCATTAGG--ATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 80

Qy       1723 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1782
Db         81 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 140

Qy       1783 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1842
Db        141 GAGACATTGAAGGGCGTCAGTCTACACCCNTTCATTTTGCAGCTGGGTATAACAGAGTGT 200

Qy       1843 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1902
Db        201 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 260
```

```
Qy       1903 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 320

Qy       1963 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 380

Qy       2023 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2082
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        381 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 440

Qy       2083 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 500

Qy       2143 AAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCA 2202
             ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        501 AAGATCTGCTTAGGGGAGATGCAGCTTAGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCA 560

Qy       2203 GAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATT 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 GAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATT 620

Qy       2263 CAACACCTTTACATTTAGCAGCTGGTTA 2290
             |||||||||||||| ||||||||||||
Db        621 CAACACCTTTACATATAGCAGCTGGTTA 648
```

RESULT 19
CD803519

```
LOCUS       CD803519                703 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  UI-M-GV0-chu-d-17-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
            IMAGE:30545656 5', mRNA sequence.
ACCESSION   CD803519
VERSION     CD803519.1  GI:32462345
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30545656"
                     /tissue_type="whole brain"
                     /dev_stage="1,5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GV0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction,ligated with
                     EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN
```

```
  Query Match             15.6%;  Score 596;  DB 14;  Length 703;
  Best Local Similarity   90.6%;  Pred. No. 8.9e-113;
  Matches  635;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy       1949 AGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACC 2008
             ||||||||||||  || ||||||||||||||| |||||  |||||  || |||||  || |||||
Db          2 AGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACC 61

Qy       2009 TTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGG 2068
             ||||||||||||  |||||||||||||||||||||||||||  ||||||||||  ||||||||||
Db         62 TTTACATGAAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGG 121

Qy       2069 TGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGG 2128
             ||||||||||| || |||||||| |||||| |||||||| ||||||||||||||||||||||
Db         122 TGCAGACCCTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGG 181

Qy       2129 AGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAA 2188
             |||||||||||||||||||| ||||| || |||||||||||||| ||||||||||||||
Db         182 AGATACAGATATTCAAGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAA 241

Qy       2189 GGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATAC 2248
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         242 GGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATAC 301

Qy       2249 CCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGC 2308
             ||||||  | ||||| |||||||||||||| ||||||||||||||||||  || ||||||||
Db         302 CCAAGGTCGGCATTCTACACCTTTACATTTGGCAGCTGGTTATAATAACTTGGAAGTTGC 361

Qy       2309 AGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCC 2368
             |||||||| || |||| |||||||||||||||||||||||||||||||||||||| |||||
Db         362 TGAGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGACAAAGGAGGACTCATTCC 421

Qy       2369 TTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAA 2428
             |||||||| |||||||| || || |||||||||||||||||| | || |||||||| ||
Db         422 TTTACATAACGCAGCATCGTATGGACATGTAGATGTAGCAGCTTTGCTGATAAAGTACAA 481

Qy       2429 TGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAA 2488
             ||| ||||||||||||||||||||||||||||||||||||| || || ||||||||||||||
Db         482 TGCGTGTGTCAATGCCACGGACAAATGGGCTTTCACACCGTTACATGAAGCAGCCCAAAA 541

Qy       2489 GGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAA 2548
             |||||| |||||||||||||||||| ||| | |||||||||||||| || ||||||||||
Db         542 GGGACGGACACAGCTTTGTGCTTTATTGTTGGCCCATGGAGCTGATCCTACTCTTAAAAA 601

Qy       2549 TCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGAC 2608
             ||| |||||||||||||||||||||||||||||||||||| |||||||||| || ||  ||||
Db         602 TCAAGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGAC 661

Qy       2609 AGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGC 2649
             ||||| |||||||| |||||||||||  | || |||||  |
Db         662 AGCAGNCATGCCCCCCTCTGCTCTGCCTACGTGCTACAACC 702
```

RESULT 20
BE220005/c

```
LOCUS       BE220005                622 bp    mRNA    linear   EST 03-JUL-2000
DEFINITION  hv65h03.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178325 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BE220005
VERSION     BE220005.1  GI:8907323
```

```
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1. .622
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3178325"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           15.5%;  Score 593.2;  DB 10;  Length 622;
  Best Local Similarity   97.1%;  Pred. No. 3.2e-112;
  Matches  604;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       3183 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3242
              | ||||||||  |||||  ||||||||||||||||||  ||||  ||||  |  | |||||  |
Db        622 GCAATCTTCACACGATACTATATTCTCAAGATTCAGCAGGTCTGTACAAGTACACTATCG 563

Qy       3243 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3302
              |||||||||||||| | ||||||| ||||||||||||||||| ||||||||||||||||
Db        562 CAAAGATACACTCACGGCAGAAAAGCAGTTTCTGAAGAAAACCCCAACCATGCCAATGAA 503

Qy       3303 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 443

Qy       3363 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 383

Qy       3423 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 323

Qy       3483 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 263

Qy       3543 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 203

Qy       3603 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 143

Qy       3663 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 83

Qy       3723 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 23

Qy       3783 CTCTACGTTTTACTCCTTTGCT 3804
              |||||||||||||||||||||||
Db         22 CTCTACGTTTTACTCCTTTGCT 1


RESULT 21
BE467629/c
LOCUS       BE467629                616 bp    mRNA    linear   EST 27-JUL-2000
DEFINITION  hz67b11.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3213021 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BE467629
VERSION     BE467629.1  GI:9513404
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 616)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 402.
FEATURES             Location/Qualifiers
     source          1. .616
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3213021"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           15.3%;  Score 584;  DB 10;  Length 616;
  Best Local Similarity   96.8%;  Pred. No. 2.6e-110;
  Matches  596;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       3189 TTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGA 3248
```

```
Db        616 TTCACAAGATCCAATATTCTCAAGATCATATAGGTTTGCAACCACGAACTATGGAATAGA 557

Qy       3249 TACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATG 3308
              ||||||||||||||| |||||||| |||||||||| || |||||||||| |||||
Db        556 TACACTCACCGGAGACAAGAAGTTTTTGAAGAAAACCCCACCCATGCCAATGACGGAATG 497

Qy       3309 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 3368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 437

Qy       3369 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 3428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 377

Qy       3429 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 317

Qy       3489 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 257

Qy       3549 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 197

Qy       3609 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        196 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 137

Qy       3669 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATA 3728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATA 77

Qy       3729 GTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTAC 3788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          76 GTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTAC 17

Qy       3789 GTTTTACTCCTTTGCT 3804
              ||||||||||||||||
Db          16 GTTTTACTCCTTTGCT 1
```

RESULT 22
BU466115

LOCUS	BU466115 793 bp mRNA linear EST 30-NOV-2002
DEFINITION	603371794F1 CSEQRBN20 Gallus gallus cDNA clone ChEST279l24 5', mRNA sequence.
ACCESSION	BU466115
VERSION	BU466115.1 GI:25959692
KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE	1 (bases 1 to 793)
AUTHORS	Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE	A Comprehensive Collection of Chicken cDNAs
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE	22335534
PUBMED	12445392
COMMENT	Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk.

FEATURES Location/Qualifiers

source	1. .793 /organism="Gallus gallus" /mol_type="mRNA" /strain="Layer and broiler" /db_xref="taxon:9031" /clone="ChEST279l24" /sex="Male and female" /tissue_type="Chondrocytes isolated from growth plate cartilage" /dev_stage="adult" /lab_host="DH10B" /clone_lib="CSEQRBN20" /note="Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: Not1; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN

Query Match 15.3%; Score 583.8; DB 13; Length 793;
Best Local Similarity 84.5%; Pred. No. 3.2e-110;
Matches 668; Conservative 0; Mismatches 122; Indels 1; Gaps 1;

```
Qy       1293 AATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATAT 1352
              || |||||||||||||||| ||||| |||||| |||| ||||||||||||||||| |||
Db          4 AACTTCAAGCATCCTCAAACGCATGAGACAGCACTGCACTGTGCTGCTGCATCTCCTTAT 63

Qy       1353 CCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAG 1412
              || || ||||| ||| ||||||||| || |||| || |||||||| ||||| |||||||||
Db         64 CCTAAGAGAAAACAAGTATGTGAATTGCTGCTGAGGAAAGGAGCTAACATTAATGAAAAG 123

Qy       1413 ACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTT 1472
              || ||||| |||||||||||||||||| |||||||| || ||||||||||||||||||||
Db        124 ACAAAAGACTTCTTGACTCCTCTGCATGTGGCATCAGAAAAAGCTCATAATGATGTTGTT 183

Qy       1473 GAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCT 1532
              |||||||| |||||||||||||||||||||||||  ||||||| ||||| ||||| |||
Db        184 GAAGTAGTTGTGAAACATGAAGCAAAGGTTAATGCATTGGATAACCTTGGCCAGACCTCT 243

Qy       1533 CTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGG 1592
              || |||||||| ||| |||||||||| || || || ||| | |  | |||| | |||
Db        244 CTTCACAGAGCGGCACATTGTGGTCACCTCCAGACTTGCAGGTTGTTGTTGAGTTCTGGA 303

Qy       1593 TGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAAT 1652
              |||||||||  |||| | ||  | |||||||||| || ||||| ||||| ||||||| |
Db        304 TGTGATCCTTCCATTGTGTCTTTGCAGGGCTTTACAGCCTTACAAATGGGGAATGAAAGT 363

Qy       1653 GTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTG 1712
              || || ||||| || ||||| |||||| |  |||||||||||| ||||| |||||||||
Db        364 GTGCAACAACTACTGCAAGAAGGTATCCCTCTAGGTAATTCAGATGCAGATAGACAATTG 423

Qy       1713 CTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGT 1772
              |||| ||||| |||||||| ||||| || ||||||||||||||||||| ||||| |||
Db        424 TTGGAGGCTGCGAAGGCTGGCGATGTGGATACTGTAAAAAAACTGTGTACGGTTCAAAGT 483

Qy       1773 GTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTAT 1832
              || |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        484 GTGAACTGCAGAGATATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGATAC 543

Qy       1833 AACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGAT 1892
              || || ||||| || || |||||||| || ||||||||||||||||||||| ||||||
Db        544 AATAGGGTGTCTGTTGTTGAATATCTTCTTCAGCATGGAGCTGATGTGCATGCAAAAGAT 603

Qy       1893 AAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAA 1952
```

```
Db        604 AAGGGAGGACTTGTTCCTTTACATAACGCTTGCTCATATGGTCACTATGAGGTTGCGGAG 663

Qy       1953 CTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTA 2012

Db        664 CTGCTTGTTAAACATGGTGCAGTTGTCAACGTAGCTGACTTGTGGAAGTTTACTCCCTTG 723

Qy       2013 CATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCA 2072

Db        724 CATGAAGCTGCAGCAAAAGGAAAATATGAGATTTGC-AACTCCTATTGCAGCATGGAGCT 782

Qy       2073 GACCCTACCAA 2083

Db        783 GACCCCAAAAA 793


RESULT 23
CF739820
LOCUS       CF739820                689 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-HD0-clf-m-21-0-UI.r1 NIH_BMAP_HD0 Mus musculus cDNA clone
            IMAGE:30611156 5', mRNA sequence.
ACCESSION   CF739820
VERSION     CF739820.1  GI:37636158
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .689
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30611156"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HD0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I;  The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH)."
ORIGIN

  Query Match           15.2%;  Score 578.8;  DB 14;  Length 689;
  Best Local Similarity   90.0%;  Pred. No. 3.3e-109;
  Matches  619;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy       2917 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 2976

Db          1 AGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCA 60

Qy       2977 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 3036

Db         61 ATGCCTATGGACATCGACACAAGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAAC 120

Qy       3037 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 3096

Db        121 AAGGTCTTAATCCATATTTAACTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATC 180

Qy       3097 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 3156

Db        181 TGTCTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGG 240

Qy       3157 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 3216

Db        241 AGCACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTC 300

Qy       3217 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 3276

Db        301 AGAAGGTTTGTAACAAGAAATTGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTG 360

Qy       3277 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 3336

Db        361 AAGAAAACCACAACCATGCAAACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATG 420

Qy       3337 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3396

Db        421 CGATTATCCATAAGGGCTTTGATGAAAGGCATGCATACATAGGTGGCATGTTTGGAGCTG 480

Qy       3397 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 3456

Db        481 GAATTTATTTTGCTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGAATTGGAGGTG 540

Qy       3457 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 3516

Db        541 GCACCGGATGTCCAATTCACAAAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCTGT 600

Qy       3517 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 3576

Db        601 TTTGTCGAGTAACCTTGGGAAAGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATT 660

Qy       3577 CTCCTCCAGGTCATCACTCAGTCACTGG 3604

Db        661 CTCCTCNCGGNCATCACTCGGTCACTGG 688


RESULT 24
BU427328
LOCUS       BU427328                796 bp    mRNA    linear   EST 29-NOV-2002
DEFINITION  603234467F1 CSEQRBN09 Gallus gallus cDNA clone ChEST235e17 5', mRNA
            sequence.
ACCESSION   BU427328
VERSION     BU427328.1  GI:25920004
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
```

(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

```
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer and broiler"
                     /db_xref="taxon:9031"
                     /clone="ChEST235e17"
                     /sex="Male and female"
                     /tissue_type="Chondrocytes isolated from growth plate
                     cartilage"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQRBN09"
                     /note="Vector: pBluescript II KS(+); Site_1: EcoRI;
                     Site_2: Not1; This normalized library was constructed from
                     1 million independent clones. cDNA synthesis was initiated
                     using an oligo(dT) primer, using methylated C in the first
                     strand synthesis reaction. Following this first strand
                     reaction, double-stranded cDNA was blunted, ligated to
                     NotI adapters, digested with EcoRI, size-selected, and
                     cloned into the NotI and EcoRI compatible sites of a
                     custom modified MCS of the pBluescript (KS+) vector. The
                     library was normalized in 2 rounds using conditions
                     adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
                     Bonaldo et al., Genome Research 6 (1996): 791, except that
                     a significantly longer reannealing hybridization was
                     used."
ORIGIN

  Query Match             15.0%;  Score 574;  DB 13;  Length 796;
  Best Local Similarity   83.5%;  Pred. No. 3.5e-108;
  Matches  663;  Conservative    0;  Mismatches  130;  Indels    1;  Gaps    1;

Qy       1293 AATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATAT 1352
             || |||||||||||||||||| ||||| |||||| |||| ||||||||||||||||||| |||
Db             4 AACTTCAAGCATCCTCAAACGCATGAGACAGCACTGCACTGTGCTGCTGCATCTCCTTAT 63

Qy       1353 CCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAG 1412
             || || ||||| ||| |||||||| || |||| || |||||||| ||||| |||||||||
Db            64 CCTAAGAGAAAACAAGTATGTGAATTGCTGCTGAGGAAAGGAGCTAACATTAATGAAAAG 123

Qy       1413 ACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTT 1472
             || ||||| |||||||||||||||||| |||||||| || |||||||||||||||||||||
Db            124 ACAAAAGACTTCTTGACTCCTCTGCATGTGGCATCAGAAAAAGCTCATAATGATGTTGTT 183

Qy       1473 GAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCT 1532
             |||||||| |||||||||||||||||||||||||  ||||||| ||||| ||||| |||
Db            184 GAAGTAGTTGTGAAACATGAAGCAAAGGTTAATGCATTGGATAACCTTGGCCAGACCTCT 243

Qy       1533 CTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGG 1592
             || ||||||||| ||| |||||||||| || || || ||| |  |  |  |||| | |||
Db            244 CTTCACAGAGCGGCACATTGTGGTCACCTCCAGACTTGCAGGTTGTTGTTGAGTTCTGGA 303

Qy       1593 TGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAAT 1652
             |||||||||  |||| | ||  | |||||||||||| || ||||| ||||| |||||||| |
Db            304 TGTGATCCTTCCATTGTGTCTTTGCAGGGCTTTACAGCCTTACAAATGGGGAATGAAAGT 363

Qy       1653 GTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTG 1712
             || || ||||| || ||||| |||||| |  |||||||||||| ||||| |||||||||
Db            364 GTGCAACAACTACTGCAAGAAGGTATCCCTCTAGGTAATTCAGATGCAGATAGACAATTG 423

Qy       1713 CTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGT 1772
              |||| ||||| |||||||| ||||| || |||||||||||||||||| ||||| |||
Db            424 TTGGAGGCTGCGAAGGCTGGCGATGTGGATACTGTAAAAAAAACTGTGTACGGTTCAAAGT 483

Qy       1773 GTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTAT 1832
             || |||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||
Db            484 GTGAACTGCAGAGATATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGATAC 543

Qy       1833 AACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGAT 1892
             || || ||||| || || |||||||| || |||||||||||||||||||||||| ||||||
Db             544 AATAGGGTGTCTGTTGTTGAATATCTTCTTCAGCATGGAGCTGATGTGCATGCAAAAGAT 603

Qy       1893 AAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAA 1952
             || ||||| ||||| ||||| || || || || || ||||| || ||||| ||||| ||
Db             604 AAGGGAGGACTTGTTCCTTTACATAACGCTTGCTCATATGGTCACTATGAGGTTGCGGAG 663

Qy       1953 CTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTA 2012
             || |||||||||||||| ||||| || || |||||||| || ||||| ||||| || ||
Db             664 CTGCTTGTTAAACATGGTGCAGTTGTCAACGTAGCTGACTTGTGGAAGTTTACTCCCTTG 723

Qy       2013 CATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCA 2072
             |||||||| ||||| |||||||||||||| |||||||||     ||  | ||||||||
Db             724 CATGAAGCTGCAGC-AAAGGAAAATATGAGATTTGCAACTCCCTATTGCAGCATGGAAGC 782

Qy       2073 GACCCTACCAAAAA 2086
                 ||  | |||||
Db             783 TGACCCCCAAAAAA 796


RESULT 25
CF735419
LOCUS       CF735419                742 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-HB0-cke-a-13-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
            IMAGE:30549036 5', mRNA sequence.
ACCESSION   CF735419
VERSION     CF735419.1  GI:37631755
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 742)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .742
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30549036"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HB0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
```

Institute of Mental Health (NIMH)."
ORIGIN

Query Match 15.0%; Score 572.8; DB 14; Length 742;
Best Local Similarity 86.5%; Pred. No. 6e-108;
Matches 642; Conservative 0; Mismatches 99; Indels 1; Gaps 1;

```
Qy       2293 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2352
Db          1 ATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGACA 60

Qy       2353 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2412
Db         61 AAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTAGATGTAGCAGCTT 120

Qy       2413 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2472
Db        121 TGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTTCACACCGTTAC 180

Qy       2473 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2532
Db        181 ATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGCCCATGGAGCTG 240

Qy       2533 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2592
Db        241 ATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 300

Qy       2593 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2652
Db        301 TCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGCTACAAACCTC 360

Qy       2653 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2712
Db        361 AAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCATTGCAGATGCTCTGTCTTCAGGTCCGT 420

Qy       2713 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2772
Db        421 CCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTGGCAGCTTCTCGG 480

Qy       2773 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2832
Db        481 AACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTTTGCANAGAAAAG 540

Qy       2833 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2892
Db        541 AGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACT 600

Qy       2893 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2952
Db        601 TGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACA 660

Qy       2953 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 3012
Db        661 AGGAACTGAAAGAGATTGG-ATCAATGCCTATGGACATCGATACANGCTGATTAAAGGAG 719

Qy       3013 TCGAGAGACTTATCTCCGGACA 3034
Db        720 TTGAAAGGCTGATCTCTGGACA 741
```

RESULT 26
CB201923
LOCUS CB201923 913 bp mRNA linear EST 05-FEB-2003
DEFINITION AGENCOURT_11289292 NIH_MGC_135 Mus musculus cDNA clone IMAGE:30140366 5', mRNA sequence.
ACCESSION CB201923
VERSION CB201923.1 GI:28236420
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 913)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: NDAM0046 row: e column: 15
High quality sequence stop: 676.
FEATURES Location/Qualifiers
source 1. .913
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30140366"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/note="Vector: pCMVSport6.1; Site_1: EcoRV; Site_2: NotI; Normalized full-length enriched library from pooled mouse embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5, and 15.5 (size selected for the 0.5-1 kb fragments) Cloned directionally, priming method: Oligo-dT. cDNA enrichment: >1k bp, Average insert size 1.6k bp. Normalization (Cot value): 7.5 kb. Priming sequence: 5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)3' Tissue contributed by, David Rowe. Library constructed by ResGen, Invitrogen Corp."
ORIGIN

Query Match 15.0%; Score 572.6; DB 14; Length 913;
Best Local Similarity 91.8%; Pred. No. 7.2e-108;
Matches 605; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```
Qy       3068 CTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGT 3127
Db          2 CTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGT 61

Qy       3128 GGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAAT 3187
Db         62 GGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGT 121

Qy       3188 CTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAG 3247
Db        122 CTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTGGGAAAG 181

Qy       3248 ATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAAT 3307
Db        182 ATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCAAACGAAAGGAT 241

Qy       3308 GCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCA 3367
Db        242 GTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTGATGAAAGGCA 301

Qy       3368 TGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAG 3427
Db        302 TGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTCCAAAAG 361

Qy       3428 CAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATC 3487
Db        362 CAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTCCAATTCACAAAGACAGATC 421

Qy       3488 TTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCT 3547
Db        422 GTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAAAGTCTTTCTT 481

Qy       3548 GCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAG 3607
Db        482 GCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGGTCACTGGCCG 541
```

```
Qy       3608 GCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTA 3667
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTA 601

Qy       3668 TCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAA 3726
             |||||| ||| ||||||||||||||||| | ||||||||||||||||||| ||||| | ||
Db        602 TCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAGGTATGGTTGATGGCTGAA 660


RESULT 27
BE910901
LOCUS       BE910901                956 bp    mRNA    linear   EST 29-SEP-2000
DEFINITION  601661978F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3962251 5',
            mRNA sequence.
ACCESSION   BE910901
VERSION     BE910901.1  GI:10407885
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 956)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9128  row: k  column: 20
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .956
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3962251"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match             14.9%;  Score 570.4;  DB 10;  Length 956;
  Best Local Similarity   85.2%;  Pred. No. 2.1e-107;
  Matches  673;  Conservative    0;  Mismatches  111;  Indels    6;  Gaps    3;

Qy       2847 GATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTT 2906
             |||||||| |||||||| ||| ||||||| |||||||||||||||| | ||||||||||||
Db          1 GATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACTTGATGGATATATTT 60

Qy       2907 GAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAG 2966
             |||||||||||||| || || ||||| ||||||| ||||| |||||||| ||||| |||
Db         61 GAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAAGGAACTGAAAGAG 120

Qy       2967 ATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATC 3026
             ||||||||||||| ||||||||| | ||||| || |||||||||| || || || |||
Db        121 ATTGGAATCAATGCCTATGGACATCGACACAAGCTGATTAAAGGAGTTGAAAGGCTGATC 180

Qy       3027 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 3086
             || |||||||||||||||| |||||||||||| | |||| ||||||||||||||||||||
Db        181 TCTGGACAACAAGGTCTTAATCCATATTTAACTCTAAACAACTCTGGTAGTGGAACAATT 240

Qy       3087 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 3146
             || |||||||||||||||||||||||||||||||| |||||||||||| |||||||||| |||
Db        241 CTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGT 300

Qy       3147 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 3206
             || || || |||||||||||||| ||||| ||||| ||  |||||||||||||||| |||
Db        301 ACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATT 360

Qy       3207 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 3266
             |||||||||||||||||||||||||||||| | |||||||||| || ||||||||||| |
Db        361 CTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTGGGAAAGATATACACACCGGAGAAAA 420

Qy       3267 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 3326
             |||||||||||||||||||||||||||| || ||| | ||| |||||||||||||||||
Db        421 GAAGTTTCTGAAGAAAACCACAACCATGCAAACGAAAGGATGTTATTTCATGGGTCTCCT 480

Qy       3327 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATG 3386
             |||||||||| |||||||| || ||||||||||||||||||| |||||||||| |||
Db        481 TTTGTGAATGCGATTATCCATAAGGGCTTTGATGAAAGGCATGCATACATAGGTGGCATG 540

Qy       3387 TTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGA 3446
             |||||||||| |||||||||||||| ||||||||||||||||||||||||||| ||||||
Db        541 TTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGA 600

Qy       3447 ATTGGAGGAGG---TACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCAC 3503
             | | | | |     | || |||||| ||||||||||||||||| |||||| |||| |||
Db        601 ACTTGGAGGTGGCACCCGGGATGTCCAATTCACAAAGACAGATCGTGTTACCTTTGTCAC 660

Qy       3504 AGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATG 3563
             |||||||||| ||| | || || |||||||| |||||||| ||| |||||  ||| ||
Db        661 AGGCAGCTGCTGTTTCGTCGAGT-ACCTTGGGCAAGTCTTTCTTGCCGTTCAACGCATTG 719

Qy       3564 AAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGG--TAGGCCCAGTGTAAATG 3621
             | |||||    |||||  || |      | ||||||    ||||       |||||
Db        720 ACAATGGACATTTCTCTCCCGGCCATCAATCGGGCCACTGGCCCGGGCCAGTGTGAAATG 779

Qy       3622 GCCTAGCATT 3631
             |||||| |||
Db        780 GCCTAGGATT 789


RESULT 28
BB339554
LOCUS       BB339554                702 bp    mRNA    linear   EST 24-OCT-2001
DEFINITION  BB339554 RIKEN full-length enriched, 10 days neonate cerebellum Mus
            musculus cDNA clone B930018O21 3', mRNA sequence.
ACCESSION   BB339554
VERSION     BB339554.2  GI:16404413
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 11, 2000 this sequence version replaced gi:9048317.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
```

URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES Location/Qualifiers
source 1. .702
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="B930018O21"
/sex="mixed"
/tissue_type="cerebellum"
/dev_stage="10 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days neonate cerebellum"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

ORIGIN

Query Match 14.9%; Score 568; DB 10; Length 702;
Best Local Similarity 89.7%; Pred. No. 5.8e-107;
Matches 621; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

```
Qy       1332 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1391
             ||||||||||||||| |||||||||||||||||| |||||||||||| |||| |||||
Db         12 TGTGCTGCTGCATCT-CATATCCCAAAAGAAAGCAGATATGTGAACTGCTGCTCAGAAAG 70

Qy       1392 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1451
             |||||||||| ||| ||||||||||||||||||||||||||||||||||||||| ||||||
Db         71 GGAGCAAACACCAACGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCTTCTGAG 130

Qy       1452 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1511
             || ||||| |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        131 AACGCTCACAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTTTG 190

Qy       1512 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1571
             |||| |||||| ||||| || ||||||||||||||| | |||||||| || || ||||||
Db        191 GATAGTCTTGGACAGACGTCATTACACAGAGCTGCACACTGTGGTCACCTGCAGACCTGC 250

Qy       1572 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1631
             ||||| |||||||||||||||||||||| ||||| |||||||||||||| || || ||
Db        251 CGCCTGCTCCTGAGCTATGGGTGTGATCCAAACATCATATCCCTTCAGGGTTTCACCGCC 310

Qy       1632 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1691
             || || |||||||||||||||| ||||| || |||||||||||    |||| ||||| |
Db        311 TTGCAAATGGGAAATGAAAATGTGCAGCAGCTGCTCCAAGAGGGCGCCTCACTAGGTCAC 370

Qy       1692 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAA 1751
             |||||||| |||||||  ||||||||||||||||||||||||| || || ||||||||||||
Db        371 TCAGAGGCGGACAGACAGCTGCTGGAAGCTGCAAAGGCTGGTGACGTTGAAACTGTAAAA 430

Qy       1752 AAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCA 1811
             ||||| |||||||||||||||||||||||||||||||||||||||||| || ||||| || |||
Db        431 AAACTCTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCAGTCAACCCCA 490

Qy       1812 CTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGA 1871
             || || |||||||||||||| |||||||||| ||||| |||||||||| || |||||||
Db        491 CTCCACTTTGCAGCTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAACATGGA 550

Qy       1872 GCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTAT 1931
             |||||||| |||||||||| |||||||| |||||||||| ||||||||||||||||||||
Db        551 GCTGATGTTCATGCTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGCATGTTCTTAT 610

Qy       1932 GGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1991
             ||||| |||||||||||||||||||||| || |||||||||||||| ||||| |||||
Db        611 GGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGAC 670

Qy       1992 TTATGGAAATTTACACCTTTACATGAAGCAGC 2023
             || ||||| || |||||||||||||||||| ||
Db        671 TTGTGGAAGTTCACACCTTTACATGAAGCTGC 702
```

RESULT 29
CF538288
LOCUS CF538288 725 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GH0-chd-p-24-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone IMAGE:30535583 5', mRNA sequence.
ACCESSION CF538288
VERSION CF538288.1 GI:34590270
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 725)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .725
/organism="Mus musculus"

/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30535583"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 14.8%; Score 564.2; DB 14; Length 725;
Best Local Similarity 86.9%; Pred. No. 3.6e-106;
Matches 631; Conservative 0; Mismatches 94; Indels 1; Gaps 1;

Qy 2458 CTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGC 2517
Db 1 CTTTCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGT 60

Qy 2518 TAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAG 2577
Db 61 TGGCCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAG 120

Qy 2578 TTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCT 2637
Db 121 TTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTA 180

Qy 2638 CTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTC 2697
Db 181 CGTGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTC 240

Qy 2698 TCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTAT 2757
Db 241 TGTCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTAT 300

Qy 2758 CTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCA 2817
Db 301 CTGGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTG 360

Qy 2818 GTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATC 2877
Db 361 GTTTGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACC 420

Qy 2878 TTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAG 2937
Db 421 TTGGACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAG 480

Qy 2938 TTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACA 2997
Db 481 TTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACA 540

Qy 2998 AACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAA 3057
Db 541 AGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAA 600

Qy 3058 CTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGT 3117
Db 601 CTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGANTAAGA-T 659

Qy 3118 TTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATG 3177
Db 660 TTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACG 719

Qy 3178 CAGGTG 3183
Db 720 CAGGCG 725

RESULT 30
BE222311/c
LOCUS BE222311 585 bp mRNA linear EST 03-JUL-2000
DEFINITION hu10d01.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3166177 3' similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION BE222311
VERSION BE222311.1 GI:8909629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 585)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 476.
FEATURES Location/Qualifiers
source 1. .585
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3166177"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Query Match 14.7%; Score 562.8; DB 10; Length 585;
Best Local Similarity 98.5%; Pred. No. 6.4e-106;
Matches 576; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy 3222 GTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAA 3281
Db 585 GTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAA 526

Qy 3282 AACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATT 3341
Db 525 AACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATT 466

Qy 3342 ATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATT 3401

```
Db        465 ATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATT 406

Qy       3402 TATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA--GGTA 3459

Db        405 TATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGNAGGAAGGNGN 346

Qy       3460 CTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTT 3519

Db        345 NNCGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTT 286

Qy       3520 GCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTC 3579

Db        285 GCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTC 226

Qy       3580 CTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAAT 3639

Db        225 CTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAAT 166

Qy       3640 ATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGA 3699

Db        165 ATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGA 106

Qy       3700 GGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCT 3759

Db        105 GGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCT 46

Qy       3760 AAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 3804

Db         45 AAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 1
```

RESULT 31
CF728039
LOCUS CF728039 681 bp mRNA linear EST 09-OCT-2003
DEFINITION UI-M-HB0-ckm-c-21-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone IMAGE:30548708 5', mRNA sequence.
ACCESSION CF728039
VERSION CF728039.1 GI:37602207
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 681)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .681
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30548708"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."
ORIGIN

Query Match 14.7%; Score 561.2; DB 14; Length 681;
Best Local Similarity 89.9%; Pred. No. 1.5e-105;
Matches 613; Conservative 0; Mismatches 68; Indels 1; Gaps 1;

```
Qy       1597 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1656

Db          1 ATCCAAACATCATATCCCTTCAGGGTTTCACCGCCTTGCAAATGGGAAATGAAAATGTGC 60

Qy       1657 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1716

Db         61 AGCAGCTGCTCCAAGAGGGCGCCTCACTAGGTCACTCAGAGGCGGACAGACAGCTGCTGG 120

Qy       1717 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1776

Db        121 AAGCTGCAAAGGCTGGTGACGTTGAAACTGTAAAAAAACTCTGTACTGTTCAGAGTGTCA 180

Qy       1777 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1836

Db        181 ACTGCAGAGACATTGAAGGACGCCAGTCAACCCCACTCCACTTTGCAGCTGGGTACAACA 240

Qy       1837 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1896

Db        241 GAGTGTCTGTGGTAGAATATCTGCTGCAACATGGAGCTGATGTTCATGCTAAAGACAAAG 300

Qy       1897 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1956

Db        301 GAGGTCTTGTACCTTTACACAATGCATGTTCTTATGGACACTATGATGTTGCAGAACTTC 360

Qy       1957 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 2016

Db        361 TTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACCTTTACATG 420

Qy       2017 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2076

Db        421 AAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGGTGCAGACC 480

Qy       2077 CTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2136

Db        481 CTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGGAGATACAG 540

Qy       2137 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2196

Db        541 ATATTCACGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTT 600

Qy       2197 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2256

Db        601 TAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATA-CCAAGGTC 659

Qy       2257 GACATTCAACACCTTTACATTT 2278

Db        660 GGCATTCTACACCTTTACATTT 681
```

RESULT 32
CF745978
LOCUS CF745978 669 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-GV0-clx-d-21-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone IMAGE:30618620 5', mRNA sequence.
ACCESSION CF745978
VERSION CF745978.1 GI:37642321
KEYWORDS EST.

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30618620"
                     /tissue_type="whole brain"
                     /dev_stage="1,5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GV0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction,ligated with
                     EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match           14.6%;  Score 557.8;  DB 14;  Length 669;
  Best Local Similarity   90.4%;  Pred. No. 7.4e-105;
  Matches  606;  Conservative    0;  Mismatches   63;  Indels    1;  Gaps    1;

Qy        989 TCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTT 1048
              ||||||||||||||||||||| |||||||||| |||||||||| |||||||||||| |
Db          1 TCATTATGAAGTAACTGAACTTCTGGTCAAGCACGGTGCCTGTGTGAATGCAATGGACCT 60

Qy       1049 GTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCT 1108
              |||||| ||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         61 GTGGCAGTTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGATTGAAGTATGTTCTCT 120

Qy       1109 TCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGA 1168
               ||| ||||||| || ||||| ||||| ||||| || ||||||||||||||||||||||
Db        121 CCTCCTAAGTTACGGGGCAGATCCAACCCTGCTGAACTGTCACAATAAAAGTGCTATAGA 180

Qy       1169 CTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTT 1228
              | |||| |||||  ||||| | || |||||||||| | ||||||||||||||||| || ||
Db        181 CCTGGCCCCCACGGCACAGCTGAAGGAAAGATTATCGTATGAATTTAAAGGCCATTCATT 240

Qy       1229 GCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAAT 1288
              |||||||||||||||||||||||||||||  ||||||||||||||| ||||||||||||
Db        241 GCTGCAAGCTGCACGAGAAGCTGATGTTACAAGAATCAAAAAACATCTTTCTCTGGAAAT 300

Qy       1289 GGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCC 1348
              |||||||||||| |||||||| |||||||||||||| ||||||||||||||||||||||

Db        301 GGTGAATTTCAAACATCCTCAGACACATGAAACAGCACTGCATTGTGCTGCTGCATCTCC 360

Qy       1349 ATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGA 1408
              |||||||||||||||||| |||||||||| |||| ||||| |||||||||| ||| ||
Db        361 ATATCCCAAAAGAAAGCAGATATGTGAACTGCTGCTCAGAAAGGGAGCAAACACCAACGA 420

Qy       1409 AAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGT 1468
              |||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||| ||||||||
Db        421 AAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCTTCTGAGAACGCTCACAATGATGT 480

Qy       1469 TGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGAC 1528
              |||||||||||||||||||||||||||||||||||||||||||| |||||| |||||| |||||
Db        481 TGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTTTGGATAGTCTTGGACAGAC 540

Qy       1529 TTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTA 1588
                ||  |||||||||||||| | |||||||| || || |||||||||||| ||||||||||
Db        541 GTCATTACACAGAGCTGCACACTGTGGTCACCTGCAGACCTGCCGCCTGCTCCTGAGCTA 600

Qy       1589 TGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGA 1648
              |||||||||| |  |||| |||||||||||||| || || || || || |||||| ||||
Db        601 TGGGTGTGAT-CANACATCATATCCCTTCAGGGTTTCACCGCCTTGCATATGGGACATGA 659

Qy       1649 AAATGTACAG 1658
              |||||| |||
Db        660 AAATGTGCAG 669


RESULT 33
BB646543
LOCUS       BB646543                659 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB646543 RIKEN full-length enriched, 10 days neonate cerebellum Mus
            musculus cDNA clone B930018O21 5', mRNA sequence.
ACCESSION   BB646543
VERSION     BB646543.1  GI:16480872
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
```

encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

```
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="B930018O21"
                     /sex="mixed"
                     /tissue_type="cerebellum"
                     /dev_stage="10 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 10 days neonate
                     cerebellum"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 100.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:
                     BamHI"
ORIGIN
```

Query Match 14.6%; Score 557.6; DB 10; Length 659;
Best Local Similarity 91.9%; Pred. No. 8.1e-105;
Matches 600; Conservative 0; Mismatches 50; Indels 3; Gaps 1;

```
Qy         69 GCCCATGGGACTGCGCCGGATCCGGTGACAGCAGGGAGCCAAGCGGCCCGGGCCCTGAGC 128
              | |||||||||| | ||||  || |||||||||| |||||||  |||||||||||||||||||
Db          7 GGCCATGGGACCGTGCCGACTCTGGTGACAGCGGGGAGCGGAGCGGCCCGGGCCCTGAGC 66

Qy        129 GCGTCTTCTCCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCGGGGCTCCTGCTCCGGTTG 188
              |||||| || |||||||||||||||||||||||||||||||||||| |||||||||| |  |||
Db         67 GCGTCTCCTGCGGGGGGCCTCGCCCTCCTGCTCGCGGGGCCCGGGCTCCTGCTGCCTTTG 126

Qy        189 CTGGCGCTGTTGCT---GGCTGTGGCGGCGGCCAGGATCATGTCGGGTCGCCGCTGCGCC 245
              |||||||| ||||    ||| | |||||||||| |||||||||| || ||||||||||||
Db        127 CTGGCGCTGCTGCTGGCGGCGGCGGCGGCGCCCGGATCATGTCTGGCCGCCGCTGCGCC 186

Qy        246 GGCGGGGGAGCGGCCTGCGCGAGCGCCGCGGCCGAGGCCGTGGAGCCGGCCGCCCGAGAG 305
              ||||||| |||||||||||||||||| |||||||||||||||||| ||||||| |||
Db        187 GGCGGGGGCGCGGCCTGCGCGAGCGCCGGGGCCGAGGCCGTGGAGCCGTCCGCCCGCGAG 246

Qy        306 CTGTTCGAGGCGTGCCGCAACGGGGACGTGGAACGAGTCAAGAGGCTGGTGACGCCTGAG 365
              |||||||||| |||||||||||||||||||| |||||||||||||||||||||||| |||
Db        247 CTGTTCGAGGCTTGCCGCAACGGGGACGTGGAGCGAGTCAAGAGGCTGGTGACGCCCGAG 306

Qy        366 AAGGTGAACAGCCGCGACACGGCGGGCAGGAAATCCACCCCGCTGCACTTCGCCGCAGGT 425
              |||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db        307 AAGGTGAACAGCCGCGACACGGCGGGCAGGAAGTCCACCCCGCTCCACTTCGCCGCAGGT 366

Qy        426 TTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCACGT 485
              ||||| ||||| || |||||||||||| | |||||||| |||||||||| ||||| |||
```

```
Db        367 TTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGCAAGCGCGT 426

Qy        486 GATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTC 545
              |||||||||| |||||||||||||||||| |||||||||||||||||||||||||||||||
Db        427 GATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTGAAGTAGTC 486

Qy        546 AATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCT 605
              ||||| ||||| | ||| |||||||||| |||||||| ||||||||||||||||||||||
Db        487 AATCTTCTTTTACAACACGGTGCAGACCCAAATGCTCGGGATAATTGGAATTATACTCCT 546

Qy        606 CTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGA 665
              |||||||||||||||||||||||| |||||||||||||||||| ||||||||||||
Db        547 CTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTACAGCATGGA 606

Qy        666 GCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGA 718
              |||||||| ||||| ||||||||||||||||||||||||||||||||||||
Db        607 GCTGAGCCCACCATNCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGA 659
```

RESULT 34
BE504175/c

```
LOCUS       BE504175                595 bp    mRNA    linear   EST 04-AUG-2000
DEFINITION  hv86c11.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3180308 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BE504175
VERSION     BE504175.1  GI:9706583
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 409.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3180308"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN
```

Query Match 14.6%; Score 557.4; DB 10; Length 595;
Best Local Similarity 99.3%; Pred. No. 8.5e-105;
Matches 591; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

```
Qy       3213 ATTCAGAAGGTTTGTAACAAG-AAACTATGGGAAAGATACACTCACCGG-AGAAAAGAAG 3270
             ||||||||||||||||||||| |||||||||||||||||||||||||| | |||||||||
Db        595 ATTCAGAAGGTTTGTAACAAGAAAACTATGGGAAAGATACACTCACCCGAAGAAAAGAAG 536

Qy       3271 TTTCTGAAGAAAACCAC-AACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3329
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        535 TTTCTGAAGAAAACCACAAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 476

Qy       3330 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3389
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 416

Qy       3390 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3449
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 356

Qy       3450 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3509
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 296

Qy       3510 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3569
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 236

Qy       3570 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3629
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 176

Qy       3630 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3689
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 116

Qy       3690 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3749
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 56

Qy       3750 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 3804
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 1
```

RESULT 35
CF530741

```
LOCUS       CF530741                640 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-FY0-cgm-g-09-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
            IMAGE:30354680 5', mRNA sequence.
ACCESSION   CF530741
VERSION     CF530741.1  GI:34582705
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30354680"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_FY0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN
```

```
  Query Match           14.5%;  Score 552;  DB 14;  Length 640;
  Best Local Similarity  91.4%;  Pred. No. 1.2e-103;
  Matches  585;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy       1711 TGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGA 1770
             |||||||||||||||||||| || ||||||||||||||||| |||||||||||||||||||
Db          1 TGCTGGAAGCTGCAAAGGCTGGTGACGTTGAAACTGTAAAAAAACTCTGTACTGTTCAGA 60

Qy       1771 GTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGT 1830
             |||||||||||||||||||||||| || ||||| || ||||| || ||||||||||||||
Db         61 GTGTCAACTGCAGAGACATTGAAGGACGCCAGTCAACCCCACTCCACTTTGCAGCTGGGT 120

Qy       1831 ATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAG 1890
             | |||||||||| ||||| |||||||||| || |||||||||||||| ||||||||||||
Db        121 ACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAACATGGAGCTGATGTTCATGCTAAAG 180

Qy       1891 ATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAG 1950
             | |||||||| |||||||||| |||||||||||||||||||||| ||||||||||||||
Db        181 ACAAAGGAGGTCTTGTACCTTTACACAATGCATGTTCTTATGGACACTATGAAGTTGCAG 240

Qy       1951 AACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTT 2010
             |||||||||| || |||||||||||||| ||||| ||||| || ||||| || |||||||
Db        241 AACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACCTT 300

Qy       2011 TACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTG 2070
             |||||||||| |||||||||||||||||||||||||||| |||||||||| ||||||||||
Db        301 TACATGAAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGGTG 360

Qy       2071 CAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAG 2130
             |||||||||| || |||||||| ||||| |||||||||| ||||||||||||||||||||
Db        361 CAGACCCTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGGAG 420

Qy       2131 ATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGG 2190
             |||||||||||||||||| ||||| || |||||||||||||| ||||||||||||||||||
Db        421 ATACAGATATTCAAGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAAGG 480

Qy       2191 GTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCC 2250
             |||||||||||||||||||| ||||||||||||||||||||||||||||| ||||||||||
Db        481 GTTGTTTAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATACCC 540

Qy       2251 AAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAG 2310
             ||||  | ||||| |||||||||||||| |||||||||||||||| || | |||||| |
Db        541 AAGGTCGGCATTCTACACCTTTACATTTGGCAGCTGGTTATAATAACTTGGTAGTTGCTG 600

Qy       2311 AGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGA 2350
             ||||||| || ||||| ||||||||||||||||||||| ||
Db        601 AGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGA 640
```

RESULT 36

```
BQ267594
LOCUS       BQ267594                553 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  ij90c10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:5778810
            5' similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BQ267594
VERSION     BQ267594.1  GI:20492659
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 429.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5778810"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Hiroshi Inoue, MD/PhD for further
                     information on this library (Metabolism Division, Permutt
                     Laboratory, Washington University School of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN
```

```
  Query Match             14.4%;  Score 549;  DB 13;  Length 553;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-103;
  Matches  549;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3260 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3319
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 64

Qy       3320 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3379
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 124

Qy       3380 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3439
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       125 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 184

Qy       3440 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3499
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 244

Qy       3500 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3559
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 304

Qy       3560 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3619
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 364

Qy       3620 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3679
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 424

Qy       3680 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3739
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 484

Qy       3740 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3799
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 544

Qy       3800 TTGCTGAAA 3808
             |||||||||
Db        545 TTGCTGAAA 553
```

```
RESULT 37
CF729155
LOCUS       CF729155                691 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  UI-M-HD0-clc-f-23-0-UI.r1 NIH_BMAP_HD0 Mus musculus cDNA clone
            IMAGE:30615214 5', mRNA sequence.
ACCESSION   CF729155
VERSION     CF729155.1  GI:37603323
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30615214"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HD0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I;  The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
```

size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 14.3%; Score 545; DB 14; Length 691;
Best Local Similarity 86.8%; Pred. No. 3.4e-102;
Matches 599; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```
Qy       2520 GCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTT 2579
Db          2 GCCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTT 61

Qy       2580 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2639
Db         62 TCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACG 121

Qy       2640 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2699
Db         122 TGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTG 181

Qy       2700 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2759
Db         182 TCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCT 241

Qy       2760 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2819
Db         242 GGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGT 301

Qy       2820 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2879
Db         302 TTGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTT 361

Qy       2880 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2939
Db         362 GGACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTT 421

Qy       2940 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 2999
Db         422 GAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACAAG 481

Qy       3000 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3059
Db         482 CTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAACT 541

Qy       3060 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3119
Db         542 CTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGANTAAGAATTT 601

Qy       3120 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3179
Db         602 CAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCA 661

Qy       3180 GGTGGAATCTTCAACAGATACAATATTCTC 3209
Db         662 GGCGGCGTCTTCAACAGATACAACATTCTC 691
```

RESULT 38
BG245969
LOCUS BG245969 1099 bp mRNA linear EST 13-FEB-2001
DEFINITION 602358786F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4487204 5', mRNA sequence.
ACCESSION BG245969
VERSION BG245969.1 GI:12755784
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1099)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10331 row: 1 column: 21
High quality sequence stop: 698.
FEATURES Location/Qualifiers
source 1. .1099
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4487204"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam1"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 14.3%; Score 544; DB 12; Length 1099;
Best Local Similarity 88.5%; Pred. No. 6.8e-102;
Matches 680; Conservative 0; Mismatches 80; Indels 8; Gaps 8;

```
Qy       2937 GTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCAC 2996
Db          1 GTTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACAC 60

Qy       2997 AAACTAATTAAAGGAGTCGAGAGACTTATCTC-CGGACAACAAGGTCTTAACCCATATTT 3055
Db         61 AAGCTGATTAAAGGAGTTGAAAGGCTGATCTCGTGGACAACAAGGTCTTAATCCATATTT 120

Qy       3056 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3115
Db        121 AACTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGA 180

Qy       3116 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3175
Db        181 ATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCA 240

Qy       3176 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 3235
Db        241 CGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 300

Qy       3236 ACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGC 3295
Db        301 ATTGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGC 360

Qy       3296 CAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTT 3355
Db        361 AAACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTT 420

Qy       3356 TGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAA 3415
Db        421 TGATGAAAGGCATGCATACATAGGTGGCATGTATGGAGCTGGAATTTATTTTGCTGAGAA 480

Qy       3416 CTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGT-GTCCAGTTC 3474
Db        481 CTCTTCCAAAAGCAATCAATATGTGTATGGAATCGGAGGTGGCACCGGATGGTCCAATTC 540
```

```
Qy       3475 ACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGG 3534
             ||||||||||||| |||||||||| |||||||||||||||   |||| || ||||||||||
Db        541 ACAAAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCT-GTTTGTCGAGTAACCTTGG 599

Qy       3535 GAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACT 3594
             |||||||||| |||||||||| ||||||||||||||||||||||||| || |||||||
Db        600 GAAAGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGGCATCACT 659

Qy       3595 CAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAG 3654
             | ||||||||  || |||||||||||| |||||||||| |||||||| ||||||||||
Db        660 CGGTCACTGGGCGG-CCAGTGTAAATGG-CTAGCATTAG-TGAATATG-TATTTACAGAG 715

Qy       3655 GAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGC 3702
             || |||||| |||| |||| || |||||||| | ||| |  | |
Db        716 GAAAACAGG-TTATACTGAAAATCTAATTACTTTACCGATGGTAGGCC 762


RESULT 39
BE504478/c
LOCUS       BE504478                562 bp    mRNA    linear   EST 04-AUG-2000
DEFINITION  hz57a02.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3212042 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BE504478
VERSION     BE504478.1  GI:9706886
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 503.
FEATURES             Location/Qualifiers
     source          1. .562
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3212042"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           14.1%;  Score 538.6;  DB 10;  Length 562;
  Best Local Similarity   97.3%;  Pred. No. 6.7e-101;
  Matches  547;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       3243 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3302
             |||||||||||||||||   |||  |  ||    |  ||| ||||||||||||||||||
Db        562 GAAAGATACACTCACCGGGAGAAAAGAAGTTTCTGAAGAAAACACAACCATGCCAATGAA 503

Qy       3303 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 443

Qy       3363 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 383

Qy       3423 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 323

Qy       3483 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 263

Qy       3543 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3602
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        262 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTNCAGGTCATCACTCAGTCACT 203

Qy       3603 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 143

Qy       3663 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 83

Qy       3723 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 23

Qy       3783 CTCTACGTTTTACTCCTTTGCT 3804
             |||||||||||||||||||||||
Db         22 CTCTACGTTTTACTCCTTTGCT 1


RESULT 40
CF538447
LOCUS       CF538447                621 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-GI0-che-m-16-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone
            IMAGE:30533583 5', mRNA sequence.
ACCESSION   CF538447
VERSION     CF538447.1  GI:34590429
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
```

/clone="IMAGE:30533583"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GI0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 14.1%; Score 537.8; DB 14; Length 621;
Best Local Similarity 91.6%; Pred. No. 1e-100;
Matches 569; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```
Qy        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGACCC 60

Qy        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781
              |||||||||||| || || |||||||| |||||||||||||||| | ||||||||||| 
Db         61 ATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGAAAGATGAACTTCTGGAAAGTGCCAG 120

Qy        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
              |||||||||||| ||||||||||||||| || || |||||||| |||||||||||||| ||
Db        121 GAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCCATGC 180

Qy        842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
              |||||||||||||||||||||||||| |||||||||||||||| |||||||||| || 
Db        181 AAGTGATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGATACAACAGAGTAAAAAT 240

Qy        902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
              |||||| |||||||||| ||||| |||||||||||||||||||||||| ||||||||
Db        241 TGTACAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATCTGGT 300

Qy        962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
              |||| |||||||||||||| ||||| |||||||||||||||||||||| ||||||||||
Db        301 ACCACTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCA 360

Qy       1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
              |||||||||||| |||||||||| ||||||| ||||||||||||||||||||||||||
Db        361 CGGTGCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAGCTTC 420

Qy       1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
              |||||||||| |||||||||||||||| ||| ||||||| || ||||| ||||| |||||
Db        421 TAAGAACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCAACCCTGCT 480

Qy       1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
              || |||||||||||||||||||||| |||| |||||  ||||| | || ||||||||
Db        481 GAACTGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAAGATT 540

Qy       1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
              | | |||||||||||||||| || ||||||||||||||||||||||||||||||   |
Db        541 ATCGTATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTACAAG 600

Qy       1262 AATCAAAAAACATCTCTCTCT 1282
              |||||||||||||| |||||
Db        601 AATCAAAAAACATCTTTCTCT 621
```

RESULT 41
CF734609

LOCUS CF734609 703 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-HB0-ckg-a-02-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone IMAGE:30546721 5', mRNA sequence.
ACCESSION CF734609
VERSION CF734609.1 GI:37630945
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 703)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30546721"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 14.1%; Score 537.6; DB 14; Length 703;
Best Local Similarity 86.7%; Pred. No. 1.2e-100;
Matches 602; Conservative 0; Mismatches 91; Indels 1; Gaps 1;

```
Qy       2591 TGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCC 2650
              |||||| || ||  |||||||||||||||||||| ||||||||||  | || ||||| ||
Db          1 TGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGCTACAAACC 60

Qy       2651 TCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCC 2710
              |||||||||| ||| |||||  |||| |||||||||||||||||||| ||||||||||
Db         61 TCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTGTCTTCAGGTCC 120

Qy       2711 ATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTC 2770
              || |||||||| ||||| |||||||||||||| || |||||||||||||| || || ||
Db        121 GTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTGGCAGCTTCTC 180

Qy       2771 AGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAA 2830
              |||||||  |||| |||||||||||| | |||| |||||| |  ||||| | | |||
Db        181 GGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTTTGCAAAGAAA 240

Qy       2831 GGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCA 2890
```

```
                |||| | ||||| | |||||||| |||||||| ||| ||||||| ||||||||||||||
Db        241 AGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCA 300

Qy       2891 CCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCA 2950
                || |||||||||||||||||||||||||| || || ||||| ||||||| ||||| ||
Db        301 CTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCA 360

Qy       2951 CAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGG 3010
                |||||| ||||| |||||||||||||||||| ||||||||| | ||||| || |||||||||
Db        361 CAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACAAGCTGATTAAAGG 420

Qy       3011 AGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTC 3070
                ||| || || || ||||| |||||||||||||||| |||||||||| | |||| |||
Db        421 AGTTGAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAACTCTAAACAACTC 480

Qy       3071 TGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGA 3130
                |||||||||||||||||| |||||||||||||||||||||||| |||||||||||||||
Db        481 TGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGTGGA 540

Qy       3131 GGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTT 3190
                |||||||||||| ||||| || || |||||||||||||| ||||| ||||| ||  ||||
Db        541 AGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGTCTT 600

Qy       3191 CAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATA 3250
                || |||||||| |||||||||||||||||||||||||||| |||| | ||||||||||
Db        601 CANCAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACANGAAATTGTGGGAAAGATA 660

Qy       3251 CACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3284
                || |||||||||||||| ||||||||||||||||
Db        661 TACACACCGGAGAAAAG-AGTTTCTGAAGAAAAC 693
```

RESULT 42
BU570530

```
LOCUS       BU570530                765 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10405516 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6622292
            5', mRNA sequence.
ACCESSION   BU570530
VERSION     BU570530.1  GI:22920830
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2875  row: f  column: 20
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6622292"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN
```

```
  Query Match           14.0%;  Score 535.8;  DB 13;  Length 765;
  Best Local Similarity   87.4%;  Pred. No. 2.9e-100;
  Matches  634;  Conservative    0;  Mismatches   29;  Indels   62;  Gaps    2;

Qy       1095 GAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAAT 1154
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAAT 62

Qy       1155 AAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTT 1214
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTT 122

Qy       1215 AAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACAT 1274
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACAT 182

Qy       1275 CTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATT------- 1327
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGATAGAG 242

Qy       1328 ------------------------------------------------------GCATTG 1333
                                                                      ||||||
Db        243 ATGCTGCAGCCTGAAACTCCAGATTATATCATTTTCCACCACTGATTTTGCAGAGCATTG 302

Qy       1334 TGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGG 1393
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGG 362

Qy       1394 AGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAA 1453
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAA 422

Qy       1454 AGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGA 1513
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGA 482

Qy       1514 TAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCG 1573
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCG 542

Qy       1574 CCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTT 1633
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 NCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTT 602

Qy       1634 ACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGT-AATT 1692
                |||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||
Db        603 ACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAAGGTATCTCATTAGGTAAATT 662

Qy       1693 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1752
                |||||||||||||||||||  |  ||  ||  ||  ||  || ||| |||
Db        663 CAGAGGCAGACAGACAATTTGCTGGGAAAGCTGCCAAAGGGCCTGGNAGAATGGTCCAAA 722

Qy       1753 AACTG 1757
                |||||
Db        723 AACTG 727
```

RESULT 43
CF726657

```
LOCUS       CF726657                679 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  UI-M-HB0-ckh-n-10-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
            IMAGE:30547809 5', mRNA sequence.
ACCESSION   CF726657
VERSION     CF726657.1  GI:37600825
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30547809"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HB0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH)."
ORIGIN

  Query Match            14.0%;  Score 535.6;  DB 14;  Length 679;
  Best Local Similarity  86.7%;  Pred. No. 3e-100;
  Matches  589;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy       2354 AGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCT 2413
             ||||||||| |||||||||||||| |||||||| || || |||||||||||||||||| |
Db             1 AGGAGGACTCATTCCTTTACATAACGCAGCATCCTATGGACATGTAGATGTAGCAGCTTT 60

Qy       2414 ACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCA 2473
              || |||||||| ||||| |||||||||||||||||||||||||||||||||| || ||
Db            61 GCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTTCACACCGTTACA 120

Qy       2474 CGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGA 2533
              |||||||||||||||||||| |||||||||||||||||| ||| | ||||||||||||||
Db           121 TGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGCCCATGGAGCTGA 180

Qy       2534 CCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGT 2593
              || |||||||||||||| |||||||||||||||||||||||||||||||||||| ||||||||
Db           181 TCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGT 240

Qy       2594 CAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCA 2653
              ||| || ||  |||||||||||||||||||| |||||||||| | || ||||| |||||
Db           241 CAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGCTACAAACCTCA 300

Qy       2654 AGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATC 2713
              |||||||| ||| |||||  |||| ||||||||||||||||||||| ||||||||||| ||
Db           301 AGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTGTCTTCAGGTCCGTC 360

Qy       2714 TAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGA 2773
              |||||||| ||||| |||||||||||||| || |||||||||||||| || || || ||
Db           361 CAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTGGCAGCTTCTCGGA 420

Qy       2774 ACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGA 2833
              ||||||  |||| |||||||||||||| | |||| |||||| |  ||||| | | ||| ||
Db           421 ACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTTTGCAAAGAAAAGA 480

Qy       2834 GGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCT 2893
              || | ||||| | |||||||| |||||||| ||| ||||||| |||||||||||||||| |
Db           481 GGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACTT 540

Qy       2894 AATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAA 2953
              |||||||||||||||||||||||||||| || || |||||  ||||||| ||||| |||||
Db           541 GATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAA 600

Qy       2954 GGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGT 3013
              ||| ||||| |||||||| |||||||| ||||||||| | ||||| || ||||||||||
Db           601 GGAACTGAAAGAGATTGGNATCAATGCCTATGGACATCGACACAAGCTGATTAAAGGAGT 660

Qy       3014 CGAGAGACTTATCTCCGGA 3032
              || || || ||||| |||
Db           661 TGAAAGGCTGATCTCTGGA 679


RESULT 44
BF585334
LOCUS       BF585334                 996 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602102008F2 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4225393 5',
            mRNA sequence.
ACCESSION   BF585334
VERSION     BF585334.1  GI:11658962
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 996)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9816  row: p  column: 02
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .996
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4225393"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Co24"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

  Query Match            14.0%;  Score 534.4;  DB 10;  Length 996;
  Best Local Similarity  83.6%;  Pred. No. 6.5e-100;
  Matches  653;  Conservative    0;  Mismatches  122;  Indels    6;  Gaps    4;

Qy       2344 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGT---AG 2400
              |||| |||||||||||||| |||||||||||||| |||||||| || || |||||   |
Db             1 CCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTATGAT 60
```

```
Qy       2401 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2460
Db         61 GTAAGCAGACTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTT 120

Qy       2461 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2520
Db         121 TCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGG 180

Qy       2521 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2580
Db         181 CCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTT 240

Qy       2581 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2640
Db         241 CAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGT 300

Qy       2641 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2700
Db         301 GCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTGT 360

Qy       2701 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2760
Db         361 CTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTG 420

Qy       2761 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2820
Db         421 GCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTT 480

Qy       2821 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2880
Db         481 TGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTG 540

Qy       2881 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2940
Db         541 GACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTG 600

Qy       2941 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3000
Db         601 AAATGGGCCACAAGGAACTGAAAGAGATTGGAATC-ATGCCTATGGACATCGACACCAAG 659

Qy       3001 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC-AAGGTCTTAACCCATATTTAACT 3059
Db         660 CTGATTAAGGAGTTGAAAGGCTGATCTCTGGACAACAAAGGTCTTAATCCATATTTTACT 719

Qy       3060 TTGAACACCTCTGGTAGTGGAACAATTCTTAT-AGATCTGTCTCCTGATGATAAAGAGTT 3118
Db         720 CTTAACAAATCTGGTAGTGGGACACATCTCATAAAATCTGTGTCCCGCGGATAANGAATC 779

Qy       3119 T 3119
Db         780 T 780
```

```
RESULT 45
CA317519
LOCUS       CA317519                769 bp    mRNA    linear   EST 09-JUL-2003
DEFINITION  UI-M-FW0-cbn-m-24-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
            IMAGE:6812065 5', mRNA sequence.
ACCESSION   CA317519
VERSION     CA317519.1  GI:24535643
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .769
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6812065"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_FW0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN
```

Query Match 14.0%; Score 533.2; DB 14; Length 769;
Best Local Similarity 91.4%; Pred. No. 1e-99;
Matches 565; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```
Qy       3109 ATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATG 3168
Db          1 ATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATG 60

Qy       3169 GAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTA 3228
Db         61 GTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTA 120

Qy       3229 ACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACA 3288
Db         121 ACAAGAAATTGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACA 180

Qy       3289 ACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACA 3348
Db         181 ACCATGCAAACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATA 240

Qy       3349 AAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTG 3408
Db         241 AGGGCTTTGATGAAAGGCATGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTG 300

Qy       3409 CTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTC 3468
Db         301 CTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTC 360

Qy       3469 CAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAA 3528
Db         361 CAATTCACAAAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAA 420

Qy       3529 CCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTC 3588
Db         421 CCTTGGGAAAGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCC 480
```

```
Qy       3589 ATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTT 3648
             ||||||| ||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCACTCGGTCACTGGCCGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTT 540

Qy       3649 ACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3708
             |||||||||||||||||||||||| ||| |||||||||||||||| | ||||||||||
Db        541 ACAGAGGAGAACAGGCTTATCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAG 600

Qy       3709 GTATGGTCGATGGATAAA 3726
             ||||||| ||||| | ||
Db        601 GTATGGTTGATGGCTGAA 618
```

Search completed: March 30, 2004, 01:40:02
Job time : 9116.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 16:39:11 ; Search time 19.7436 Seconds
(without alignments)
2876.304 Million cell updates/sec

Title: US-09-843-159B-3
Perfect score: 5769
Sequence: 1 GFGRKDVVEYLLQNGASVQA.....AYPEYLITYQIMRPEGMVDG 1100

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5769	100.0	1100	4	US-09-696-668-3	Sequence 3, Appli
2	5766	99.9	1166	4	US-09-972-115A-6	Sequence 6, Appli
3	5766	99.9	1240	4	US-09-696-668-4	Sequence 4, Appli
4	5726.5	99.3	1100	4	US-09-427-154-2	Sequence 2, Appli
5	5717	99.1	1166	4	US-09-350-982C-5	Sequence 5, Appli
6	4923.5	85.3	1333	4	US-09-972-115A-2	Sequence 2, Appli
7	4889.5	84.8	1327	3	US-09-196-387-2	Sequence 2, Appli
8	4889.5	84.8	1327	4	US-09-841-835-2	Sequence 2, Appli
9	4889.5	84.8	1327	4	US-09-972-115A-8	Sequence 8, Appli
10	4847	84.0	1267	4	US-09-972-115A-4	Sequence 4, Appli
11	3093	53.6	949	3	US-09-196-387-10	Sequence 10, Appl
12	3093	53.6	949	4	US-09-841-835-10	Sequence 10, Appl
13	1918.5	33.3	673	3	US-09-196-387-8	Sequence 8, Appli
14	1918.5	33.3	673	4	US-09-841-835-8	Sequence 8, Appli
15	874	15.1	160	4	US-09-972-115A-20	Sequence 20, Appl
16	860.5	14.9	1839	2	US-09-172-977-4	Sequence 4, Appli
17	860.5	14.9	1839	4	US-09-404-108-4	Sequence 4, Appli
18	841.5	14.6	843	2	US-09-172-977-3	Sequence 3, Appli
19	841.5	14.6	843	4	US-09-404-108-3	Sequence 3, Appli
20	821	14.2	1745	2	US-09-031-485-33	Sequence 33, Appl
21	821	14.2	1745	2	US-08-847-429A-33	Sequence 33, Appl
22	821	14.2	1745	3	US-09-065-474-33	Sequence 33, Appl
23	821	14.2	1745	4	US-09-557-034-33	Sequence 33, Appl
24	681	11.8	994	4	US-10-164-595-38	Sequence 38, Appl
25	629.5	10.9	1088	3	US-09-082-059-2	Sequence 2, Appli
26	601	10.4	683	4	US-10-164-595-71	Sequence 71, Appl
27	583	10.1	1619	4	US-09-392-812A-4	Sequence 4, Appli
28	541	9.4	1704	4	US-09-392-812A-2	Sequence 2, Appli
29	536.5	9.3	1709	4	US-09-392-812A-6	Sequence 6, Appli
30	513	8.9	352	3	US-09-065-474-139	Sequence 139, App
31	513	8.9	352	4	US-09-557-034-139	Sequence 139, App
32	440	7.6	90	4	US-09-972-115A-22	Sequence 22, Appl
33	421.5	7.3	303	2	US-09-031-485-23	Sequence 23, Appl
34	421.5	7.3	303	2	US-08-847-429A-23	Sequence 23, Appl
35	421.5	7.3	303	3	US-09-065-474-23	Sequence 23, Appl
36	421.5	7.3	303	4	US-09-557-034-23	Sequence 23, Appl
37	409	7.1	302	2	US-09-031-485-38	Sequence 38, Appl
38	409	7.1	302	2	US-08-847-429A-38	Sequence 38, Appl
39	409	7.1	302	3	US-09-065-474-38	Sequence 38, Appl
40	409	7.1	302	4	US-09-557-034-38	Sequence 38, Appl
41	392	6.8	348	2	US-09-031-485-28	Sequence 28, Appl
42	392	6.8	348	2	US-08-847-429A-28	Sequence 28, Appl
43	392	6.8	348	3	US-09-065-474-28	Sequence 28, Appl
44	392	6.8	348	4	US-09-557-034-28	Sequence 28, Appl
45	390.5	6.8	786	4	US-09-509-802-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-696-668-3
; Sequence 3, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Chan, Eva
; APPLICANT: Xu, Xiang
; APPLICANT: Huang, Betty
; APPLICANT: Ossovskaya, Valeria
; TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS (
; FILE REFERENCE: A-68292-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/696,668
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/427,154
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1100
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-696-668-3

Query Match 100.0%; Score 5769; DB 4; Length 1100;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1081 AYPEYLITYQIMRPEGMVDG 1100
```

RESULT 2
US-09-972-115A-6
; Sequence 6, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-115A-6

Query Match 99.9%; Score 5766; DB 4; Length 1166;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1099; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166


RESULT 3
US-09-696-668-4
; Sequence 4, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
;  APPLICANT: Luo, Ying
;  APPLICANT:  Chan, Eva
;  APPLICANT:  Xu, Xiang
;  APPLICANT:  Huang, Betty
;  APPLICANT:  Ossovskaya, Valeria
;  TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS
;  FILE REFERENCE: A-68292-1/RMS/DHR
;  CURRENT APPLICATION NUMBER: US/09/696,668
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 09/427,154
;  PRIOR FILING DATE: 1999-10-25
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 1240
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-696-668-4

  Query Match             99.9%;  Score 5766;  DB 4;  Length 1240;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1099;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        141 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 200

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 260

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 320

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 380

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 440

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 500

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 560

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        561 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 620

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 680

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 740

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        741 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 800

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        801 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 860

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        861 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 920

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        921 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 980

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        981 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 1040

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1041 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1100

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1101 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1160

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1161 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1220

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1221 AYPEYLITYQIMRPEGMVDG 1240


RESULT 4
US-09-427-154-2
; Sequence 2, Application US/09427154
; Patent No. 6589725
; GENERAL INFORMATION:
;  APPLICANT:  Luo, Ying
;  APPLICANT:  Chan, Eva
;  APPLICANT:  Xu, Xiang
;  APPLICANT:  Huang, Betty
;  TITLE OF INVENTION:  TANKYRASE H, COMPOSTIONS INVOLVED IN THE CELL CYCLE AND
;  TITLE OF INVENTION:  METHODS OF USE
;  FILE REFERENCE:  A-68292-DJB/RMS/DAV
;  CURRENT APPLICATION NUMBER: US/09/427,154
;  CURRENT FILING DATE:  1999-10-25
;  NUMBER OF SEQ ID NOS:  9
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1100
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-427-154-2

  Query Match             99.3%;  Score 5726.5;  DB 4;  Length 1100;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 1094;  Conservative    2;  Mismatches    3;  Indels    1;  Gaps    1;
```

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db          1 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         61 PLHEAAIKGKIDVCIVLLQHGAEPTILNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIHYLLRGDAALLDAAKKGCLARVKKLSS 600

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SFSELSSLVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 -SCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1079

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
```

```
                ||||||||||||||||||||
Db       1080 AYPEYLITYQIMRPEGMVDG 1099
```

```
RESULT 5
US-09-350-982C-5
; Sequence 5, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
;  APPLICANT:  Berthelsen, Jens
;  APPLICANT:  Toma, Salvatore
;  APPLICANT:  Isacchi, Antonella
;  TITLE OF INVENTION:  Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Relat
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  PHRM-0043
;  CURRENT APPLICATION NUMBER: US/09/350,982C
;  CURRENT FILING DATE:  1999-07-09
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 5
;  LENGTH: 1166
;  TYPE: PRT
;  ORGANISM: Artificial
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Xaa is any amino acid
;  NAME/KEY: misc_feature
;  LOCATION: (1102)..(1102)
;  OTHER INFORMATION: n is any nucleic acid
;  NAME/KEY: misc_feature
;  LOCATION: (2650)..(2650)
;  OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-5

  Query Match             99.1%;  Score 5717;  DB 4;  Length 1166;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 1091;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;
```

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db         67 GFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 126

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 186

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 246

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 306

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 366

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              | ||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||
Db        367 AXHCAAASPYPKRKQICELLLRKGAXINEKTKEFLTPLHVASEKAHNDXVEVVVKHEAKV 426

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 486

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 546

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 QHGADVHAKDKXXLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 606
```

```
Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        607 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKSCLARVKKLSS 666

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 726

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        727 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALXLAHGADPTLKNQEGQTPLDLVS 786

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 846

Qy        781 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVE 840
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        847 SFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRXLGLEHLMDIFEREQITLDVLVE 906

Qy        841 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 MGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQ 966

Qy        901 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 SVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANE 1026

Qy        961 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 RMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKD 1086

Qy       1021 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 RSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ 1146

Qy       1081 AYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||
Db       1147 AYPEYLITYQIMRPEGMVDG 1166


RESULT 6
US-09-972-115A-2
; Sequence 2, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1333
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: The 'Xaa' at location 1 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (2)..(2)
;   OTHER INFORMATION: The 'Xaa' at location 2 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (3)..(3)
;   OTHER INFORMATION: The 'Xaa' at location 3 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (4)..(4)
;   OTHER INFORMATION: The 'Xaa' at location 4 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: The 'Xaa' at location 5 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (6)..(6)
;   OTHER INFORMATION: The 'Xaa' at location 6 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: The 'Xaa' at location 7 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: The 'Xaa' at location 8 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (9)..(9)
;   OTHER INFORMATION: The 'Xaa' at location 9 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, T
;   OTHER INFORMATION: yr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (10)..(10)
;   OTHER INFORMATION: The 'Xaa' at location 10 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (11)..(11)
;   OTHER INFORMATION: The 'Xaa' at location 11 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (12)..(12)
;   OTHER INFORMATION: The 'Xaa' at location 12 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (13)..(13)
;   OTHER INFORMATION: The 'Xaa' at location 13 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (14)..(14)
;   OTHER INFORMATION: The 'Xaa' at location 14 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (15)..(15)
;   OTHER INFORMATION: The 'Xaa' at location 15 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (16)..(16)
;   OTHER INFORMATION: The 'Xaa' at location 16 stands for Lys, Asn, Arg, Ser, Thr, Ile,
;   OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
```

; NAME/KEY: misc_feature
; LOCATION: (17)..(17)
; OTHER INFORMATION: The 'Xaa' at location 17 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: The 'Xaa' at location 18 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (19)..(19)
; OTHER INFORMATION: The 'Xaa' at location 19 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: The 'Xaa' at location 20 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: The 'Xaa' at location 21 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: The 'Xaa' at location 22 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: The 'Xaa' at location 23 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: The 'Xaa' at location 24 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: The 'Xaa' at location 25 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (26)..(26)
; OTHER INFORMATION: The 'Xaa' at location 26 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: The 'Xaa' at location 27 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (28)..(28)
; OTHER INFORMATION: The 'Xaa' at location 28 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: The 'Xaa' at location 29 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (30)..(30)
; OTHER INFORMATION: The 'Xaa' at location 30 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: The 'Xaa' at location 31 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: The 'Xaa' at location 32 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (33)..(33)
; OTHER INFORMATION: The 'Xaa' at location 33 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (34)..(34)
; OTHER INFORMATION: The 'Xaa' at location 34 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (35)..(35)
; OTHER INFORMATION: The 'Xaa' at location 35 stands for Lys, Asn, Arg, Ser, Thr, Ile,
; OTHER INFORMATION: Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (36)..(36)
; OTHER INFORMATION: The 'Xaa' at location 36 stands for Lys, Asn, Arg, Ser, Thr, Ile,

Query Match 85.3%; Score 4923.5; DB 4; Length 1333;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 962; Conservative 7; Mismatches 131; Indels 9; Gaps 2;

```
Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||:  :|||||||:|   :|||||| |||||||: :||||||||||||||||||||
Db        225 GFGRKNXXKYLLQNGANXQXLYNGGLIPLHXACSFGHAKXINLLLRHGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        705 QHGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||||||||||||||| ||||||| ||||| ||||| ||||||||||||||||||||
Db        765 ICKLLLQHGADPTKKNRDGNTLLDLVKDGXTDIQDXLRGDAVLLDAAKKGCLARVKKXXF 824

Qy        601 PDNVNCRDTQGRHSTPLHLA--AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGH 658
              |||||||||||||||||||
Db        825 PDNVNCRDTQGRHSTPLHLAGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 884
```

```
Qy        659 VDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDL 718

Db        885 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 944

Qy        719 VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNL 778

Db        945 XXXXXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNL 1004

Qy        779 SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838

Db       1005 SGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVL 1064

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898

Db       1065 VEMGHKELKEIXINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 1124

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK-------KLWERYTHRRKEVS 951

Db       1125 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNXXXXXAKIRHEERYTHRRKEVS 1184

Qy        952 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1011

Db       1185 EENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGG 1244

Qy       1012 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1071

Db       1245 GTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALA 1304

Qy       1072 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100

Db       1305 EYVIYRGEQAYPEYLITYQIMRPEGMVDG 1333
```

```
RESULT 7
US-09-196-387-2
; Sequence 2, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-196-387-2

  Query Match             84.8%;  Score 4889.5;  DB 3;  Length 1327;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  910;  Conservative   98;  Mismatches   80;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838

Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898

Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
```

```
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018

Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078

Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095

Db       1298 EQAYPEYLITYQIMKPE 1314


RESULT 8
US-09-841-835-2
; Sequence 2, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1327 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-841-835-2

  Query Match             84.8%;  Score 4889.5;  DB 4;  Length 1327;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  910;  Conservative   98;  Mismatches   80;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60

Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120

Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180

Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240

Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300

Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360

Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420

Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480

Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540

Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600

Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660

Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720

Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780

Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838

Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898

Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958

Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018

Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078

Db       1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy       1079 EQAYPEYLITYQIMRPE 1095

Db       1298 EQAYPEYLITYQIMKPE 1314


RESULT 9
US-09-972-115A-8
```

```
; Sequence 8, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 1327
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-8

  Query Match             84.8%;  Score 4889.5;  DB 4;  Length 1327;
  Best Local Similarity   83.0%;  Pred. No. 0;
  Matches  910;  Conservative   98;  Mismatches   80;  Indels    9;  Gaps    2;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| |||||||||||||||||||||:||| |||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||||:| |||||::|||||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||||:|||||||||||||||| ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||:|||||||||||||||||||||||||||||||||||:||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:| |
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSESTP 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLL 480
              :  |: | :||||:||||:||||:||: |:|||||:||| ||||||||||||||||||
Db        645 IRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYNRVSVVEYLL 704

Qy        481 QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYE 540
              |||||||||||||||||||||||||||||:||| ||||||||||||||||||||||||
Db        705 HHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVADLWKFTPLHEAAAKGKYE 764

Qy        541 ICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSS 600
              |||||:|||||||||||||||||:|||||||:|||||||||||||||||:|| :
Db        765 ICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLKGDAALLDAAKKGCLARVQKLCT 824

Qy        601 PDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVD 660
              |:|:||||||||:|||||||||||||||||||:|||||||||||||||||||||||||||
Db        825 PENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADVNAQDKGGLIPLHNAASYGHVD 884

Qy        661 VAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVS 720
              :|||||||| |||||||||||||||||||||||||||||||:|||||||||| :
Db        885 IAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTMKNQEGQTPLDLAT 944

Qy        721 ADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSG 780
              |||: |||  |||| |||:|:|||           : :| |  |:|| ||||||:|||:|
Db        945 ADDIRALLIDAMPPEALPTCFKPQAT-------VVSASLISPASTPSCLSAASSIDNLTG 997

Qy        781 SFSELSSVVSSSGTEGASSLEKK--EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVL 838
              :||:   :|:  :||:  |:|  || |:| :|:||:::|||||| |||| ||||||||
Db        998 PLAELAVGGASNAGDGAAGTERKEGEVAGLDMNISQFLKSLGLEHLRDIFETEQITLDVL 1057

Qy        839 VEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKE 898
              :|||:|||||||||||||||||||||: |||| |||||  :    ||||:||:|:|||
Db       1058 ADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNPYLTFHCVNQGTILLDLAPEDKE 1117

Qy        899 FQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHA 958
              :||||||||||:|||||||:|||||||||:::|||| |||| ||: ||:||||||||||
Db       1118 YQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKVVNKKLRERFCHRQKEVSEENHNHH 1177

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              ||||||||||:|||||||||||||||||||||||||||||||||||||||||||||| |
Db       1178 NERMLFHGSPFINAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTH 1237

Qy        1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
               ||||||||||:|||||||||||||| |||||:|||||| |||||||| |||||||
Db        1238 KDRSCYICHRQMLFCRVTLGKSFLQFSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRG 1297

Qy        1079 EQAYPEYLITYQIMRPE 1095
               |||||||||||||:||
Db        1298 EQAYPEYLITYQIMKPE 1314


RESULT 10
US-09-972-115A-4
; Sequence 4, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 1267
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: The 'Xaa' at location 1 stands for Thr, Ala, Pro, or Ser.
;   NAME/KEY: misc_feature
;   LOCATION: (42)..(42)
;   OTHER INFORMATION: The 'Xaa' at location 42 stands for Lys, Arg, Thr, or Met.
;   NAME/KEY: misc_feature
;   LOCATION: (98)..(98)
;   OTHER INFORMATION: The 'Xaa' at location 98 stands for Lys, Arg, Thr, or Met.
;   NAME/KEY: misc_feature
;   LOCATION: (110)..(110)
;   OTHER INFORMATION: The 'Xaa' at location 110 stands for Thr, Ala, Pro, or Ser.
;   NAME/KEY: misc_feature
;   LOCATION: (114)..(114)
;   OTHER INFORMATION: The 'Xaa' at location 114 stands for Asn, Ser, Thr, or Ile.
;   NAME/KEY: misc_feature
;   LOCATION: (126)..(126)
;   OTHER INFORMATION: The 'Xaa' at location 126 stands for Lys, or Asn.
```

; NAME/KEY: misc_feature
; LOCATION: (139)..(139)
; OTHER INFORMATION: The 'Xaa' at location 139 stands for Ile, Val, Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (144)..(144)
; OTHER INFORMATION: The 'Xaa' at location 144 stands for Thr.
; NAME/KEY: misc_feature
; LOCATION: (153)..(153)
; OTHER INFORMATION: The 'Xaa' at location 153 stands for Thr.
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: The 'Xaa' at location 166 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (173)..(173)
; OTHER INFORMATION: The 'Xaa' at location 173 stands for Ile, Val, Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (174)..(174)
; OTHER INFORMATION: The 'Xaa' at location 174 stands for Glu, or Lys.
; NAME/KEY: misc_feature
; LOCATION: (183)..(183)
; OTHER INFORMATION: The 'Xaa' at location 183 stands for Val.
; NAME/KEY: misc_feature
; LOCATION: (185)..(185)
; OTHER INFORMATION: The 'Xaa' at location 185 stands for Thr, Ala, Pro, or Ser.
; NAME/KEY: misc_feature
; LOCATION: (188)..(188)
; OTHER INFORMATION: The 'Xaa' at location 188 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (202)..(202)
; OTHER INFORMATION: The 'Xaa' at location 202 stands for His, or Pro.
; NAME/KEY: misc_feature
; LOCATION: (204)..(204)
; OTHER INFORMATION: The 'Xaa' at location 204 stands for Glu, or Lys.
; NAME/KEY: misc_feature
; LOCATION: (205)..(205)
; OTHER INFORMATION: The 'Xaa' at location 205 stands for Asn, Ser, Thr, or Ile.
; NAME/KEY: misc_feature
; LOCATION: (207)..(207)
; OTHER INFORMATION: The 'Xaa' at location 207 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (211)..(211)
; OTHER INFORMATION: The 'Xaa' at location 211 stands for Gln, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (213)..(213)
; OTHER INFORMATION: The 'Xaa' at location 213 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (215)..(215)
; OTHER INFORMATION: The 'Xaa' at location 215 stands for Lys, Asn, Glu, Asp, Gln, His
; OTHER INFORMATION: , a stop codon, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (227)..(227)
; OTHER INFORMATION: The 'Xaa' at location 227 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (228)..(228)
; OTHER INFORMATION: The 'Xaa' at location 228 stands for Asn, Asp, His, or Tyr.
; NAME/KEY: misc_feature
; LOCATION: (237)..(237)
; OTHER INFORMATION: The 'Xaa' at location 237 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (238)..(238)
; OTHER INFORMATION: The 'Xaa' at location 238 stands for Asn, Ser, Thr, Ile, Asp, Gly
; OTHER INFORMATION: , Ala, Val, His, Arg, Pro, Leu, Tyr, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (667)..(667)
; OTHER INFORMATION: The 'Xaa' at location 667 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (725)..(725)
; OTHER INFORMATION: The 'Xaa' at location 725 stands for Lys, or Asn.
; NAME/KEY: misc_feature
; LOCATION: (727)..(727)
; OTHER INFORMATION: The 'Xaa' at location 727 stands for His, Arg, Pro, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (735)..(735)
; OTHER INFORMATION: The 'Xaa' at location 735 stands for Glu, or Asp.
; NAME/KEY: misc_feature
; LOCATION: (741)..(741)
; OTHER INFORMATION: The 'Xaa' at location 741 stands for Met, Val, or Leu.
; NAME/KEY: misc_feature
; LOCATION: (763)..(763)
; OTHER INFORMATION: The 'Xaa' at location 763 stands for Leu, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (764)..(764)
; OTHER INFORMATION: The 'Xaa' at location 764 stands for Tyr, Cys, Ser, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (787)..(787)
; OTHER INFORMATION: The 'Xaa' at location 787 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (788)..(788)
; OTHER INFORMATION: The 'Xaa' at location 788 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (789)..(789)
; OTHER INFORMATION: The 'Xaa' at location 789 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (790)..(790)
; OTHER INFORMATION: The 'Xaa' at location 790 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (791)..(791)
; OTHER INFORMATION: The 'Xaa' at location 791 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (792)..(792)
; OTHER INFORMATION: The 'Xaa' at location 792 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (793)..(793)
; OTHER INFORMATION: The 'Xaa' at location 793 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (794)..(794)
; OTHER INFORMATION: The 'Xaa' at location 794 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (795)..(795)
; OTHER INFORMATION: The 'Xaa' at location 795 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (796)..(796)
; OTHER INFORMATION: The 'Xaa' at location 796 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (797)..(797)
; OTHER INFORMATION: The 'Xaa' at location 797 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (798)..(798)
; OTHER INFORMATION: The 'Xaa' at location 798 stands for Lys, Asn, Arg, Ser, Thr, Ile
; OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
; OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
; NAME/KEY: misc_feature
; LOCATION: (799)..(799)
; OTHER INFORMATION: The 'Xaa' at location 799 stands for Lys, Asn, Arg, Ser, Thr, Ile

```
;   OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (800)..(800)
;   OTHER INFORMATION: The 'Xaa' at location 800 stands for Lys, Asn, Arg, Ser, Thr, Ile
;   OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (801)..(801)
;   OTHER INFORMATION: The 'Xaa' at location 801 stands for Lys, Asn, Arg, Ser, Thr, Ile
;   OTHER INFORMATION: , Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon,
;   OTHER INFORMATION: Tyr, Trp, Cys, or Phe.
;   NAME/KEY: misc_feature
;   LOCATION: (802)..(802)

  Query Match            84.0%;  Score 4847;  DB 4;  Length 1267;
  Best Local Similarity  86.1%;  Pred. No. 0;
  Matches  949;  Conservative    6;  Mismatches  143;  Indels    4;  Gaps    2;

Qy          2 FGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTP 61
              |||||:  ||| |||: | || |||||||||||||| |  : ||| | | ||||||||||
Db        167 FGRKDLXXYLLTNGANXQXRDXGGLIPLHNACSFGXAXXIXLLLXHXAXPNARDNWNYTP 226

Qy         62 LHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 121
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 XXEAAIKGKIXXCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARS 286

Qy        122 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVP 346

Qy        182 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 LHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLN 406

Qy        242 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETA 466

Qy        302 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 LHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVN 526

Qy        362 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 ALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISL 586

Qy        422 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQ 646

Qy        482 HGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 541
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        647 HGADVHAKDKGGLVPLHNACXYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEI 706

Qy        542 CKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSP 601
              |||||||||||||| : |   |||||| |||| ||||| |||||||||||||||    |
Db        707 CKLLLQHGADPTKKKKKGXIXLDLVKDGXTDIQDXLRGDAVLLDAAKKGCLARVKKXXFP 766

Qy        602 DNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDV 661
              |||||||||||||||||| ||
Db        767 DNVNCRDTQGRHSTPLHL-AGXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 825

Qy        662 AALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSA 721

Db        826 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 885

Qy        722 DDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 781
                    ||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db        886 XXXXXXXTAAMPPSVLPSCNKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGS 945

Qy        782 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEM 841
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||

Db        946 FSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFXREQITLDVLVEM 1005

Qy        842 GHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 901
              |||||||| |||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       1006 GHKELKEIXINAYGHRHKLIKSFERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQS 1065

Qy        902 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNK---KLWERYTHRRKEVSEENHNHA 958
              |||||||||||||||||||||||||||||||||||||:   :  ||||||||||||||||||
Db       1066 VEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNRAKIRHEERYTHRRKEVSEENHNHA 1125

Qy        959 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 NERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVH 1185

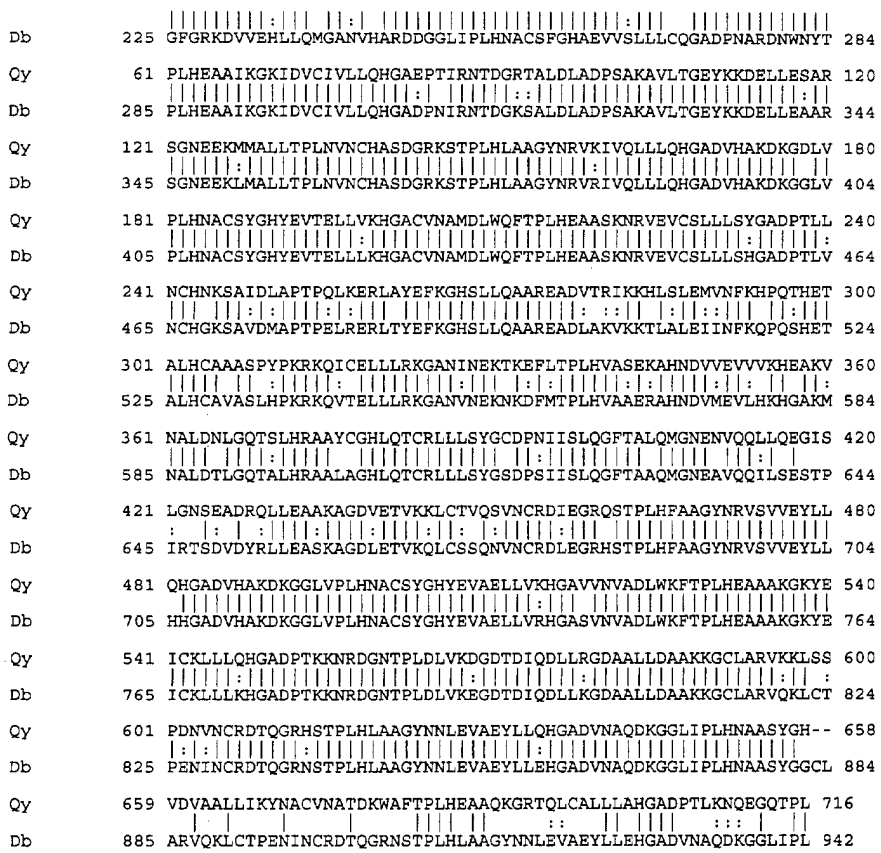
Qy       1019 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 KDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRG 1245

Qy       1079 EQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||
Db       1246 EQAYPEYLITYQIMRPEGMVDG 1267


RESULT 11
US-09-196-387-10
; Sequence 10, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 949 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-196-387-10

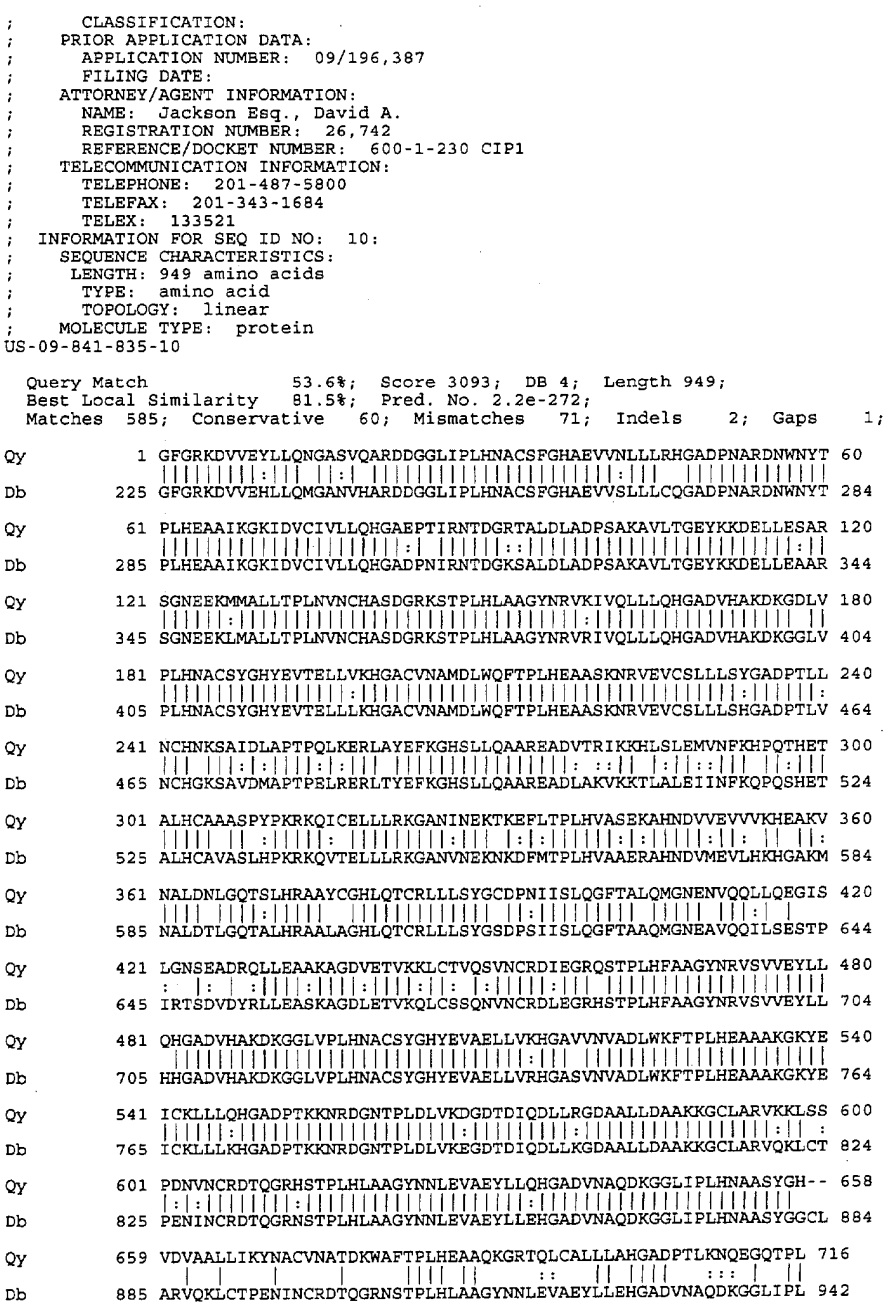
  Query Match            53.6%;  Score 3093;  DB 3;  Length 949;
  Best Local Similarity  81.5%;  Pred. No. 2.2e-272;
  Matches  585;  Conservative   60;  Mismatches   71;  Indels    2;  Gaps    1;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
```



RESULT 12
US-09-841-835-10
; Sequence 10, Application US/09841835
; Patent No. 6506587
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 949 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-841-835-10

Query Match 53.6%; Score 3093; DB 4; Length 949;
Best Local Similarity 81.5%; Pred. No. 2.2e-272;
Matches 585; Conservative 60; Mismatches 71; Indels 2; Gaps 1;



```
RESULT 13
US-09-196-387-8
; Sequence 8, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 673 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-196-387-8

  Query Match             33.3%;  Score 1918.5;  DB 3;  Length 673;
  Best Local Similarity   80.9%;  Pred. No. 1.3e-165;
  Matches  364;  Conservative   42;  Mismatches   43;  Indels    1;  Gaps    1;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:|||  ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||:| |||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||:||||||||||||||||||  ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464

Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| |||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||||| ||:|||||||| ||||| |||:|
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYG 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQS 450
                  |    |   ||: |: | |::: :  |
Db        645 SDPSIISLQGFTAAQMGN-EAVQQILSGHS 673


RESULT 14
US-09-841-835-8
; Sequence 8, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;     LENGTH: 673 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-841-835-8

  Query Match             33.3%;  Score 1918.5;  DB 4;  Length 673;
  Best Local Similarity   80.9%;  Pred. No. 1.3e-165;
  Matches  364;  Conservative   42;  Mismatches   43;  Indels    1;  Gaps    1;

Qy          1 GFGRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYT 60
              |||||||||:||| ||:| ||||||||||||||||||||:|||  ||||||||||||
Db        225 GFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNYT 284

Qy         61 PLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR 120
              |||||||||||||||||||:| |||||::|||||||||||||||||||||:||
Db        285 PLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLEAAR 344

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLV 180
              ||||||:|||||||||||||||||||||||||:||||||||||||||||||  ||
Db        345 SGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKDKGGLV 404

Qy        181 PLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLL 240
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||:
Db        405 PLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSHGADPTLV 464
```

```
Qy        241 NCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHET 300
              ||| |||:|:||||:|:||| ||||||||||||||||: ::|| |:||::||| ||:|||
Db        465 NCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIINFKQPQSHET 524

Qy        301 ALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              ||||| || :|||||: |||||||||:||| |:|:||||||:|:|||||:||: || ||:
Db        525 ALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAERAHNDVMEVLHKHGAKM 584

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGIS 420
              |||| ||||:|||||  |||||||||||| ||:|||||||| ||||| |||:|
Db        585 NALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMGNEAVQQILSVSYG 644

Qy        421 LGNSEADRQLLEAAKAGDVETVKKLCTVQS 450
                  |    |  ||: |: | |::: :  |
Db        645 SDPSIISLQGFTAAQMGN-EAVQQILSGHS 673


RESULT 15
US-09-972-115A-20
; Sequence 20, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 160
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-20

  Query Match             15.1%;  Score 874;  DB 4;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-71;
  Matches  160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        941 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSS 60

Qy       1001 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVT 120

Qy       1061 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 1100
              ||||||||||||||||||||||||||||||||||||||||
Db        121 GRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG 160


RESULT 16
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Yue, Henry
;  TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
;  FILE REFERENCE: PF-0615 US
;  CURRENT APPLICATION NUMBER: US/09/172,977
;  CURRENT FILING DATE: 1998-10-14
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PERL Program
; SEQ ID NO 4
;   LENGTH: 1839
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: g29491
US-09-172-977-4

  Query Match             14.9%;  Score 860.5;  DB 2;  Length 1839;
  Best Local Similarity   29.8%;  Pred. No. 2.1e-68;
  Matches  283;  Conservative  127;  Mismatches  386;  Indels  153;  Gaps   21;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  :|: ||  |:|| :   |  || |  | |||| :|:: ||: ||:    :|||
Db         75 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPL 134

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL--- 115
              : || :  |||   ||::||  :    || |    ||      | |:|    | ::
Db        135 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLP 194

Qy        116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS----------TPLHLAAGYNRVKIVQL 163
               |  |   :: |  |||     | | :| :           ||||:|| |  | :  |
Db        195 ALHIAARKDDTKSAALLLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 251

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
              ||  || |    :  : ||| |   |:  : :||:  |  ::|       |||| ||
Db        252 LLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 311

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 283
               :|  |||  || | |    |                  | | |  | : |     |
Db        312 DQVVELLLERGA-PLLARTKN-------------------GLSPLHMAAQGDHVECVK 349

Qy        284 HLSLEMVNFKHP-----QTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPL 338
              ||    :| |     : |||| ||    :   :: :||| | || | :     |||
Db        350 HL----LQHKAPVDDVTLDYLTALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPL 402

Qy        339 HVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISL 398
              |:| :|    |:|::||: | : |:   | | :| ||: |||    |||  |  |:: ::
Db        403 HIACKKNRIKVMELLVKYGASIQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNI 462

Qy        399 QGFTALQMGN-----ENVQQLLQEGI----------------SLGNSEADRQLLE---- 432
              :| ||| |       | |: ||: |                  || :|  : ||:
Db        463 RGETALHMAARAGQVEVVRCLLRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAH 522

Qy        433 --------------AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
                           :| | |:    |    : :   : :  |||| || |  : | :
Db        523 PDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKL 581

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
              |||  |   :  | || ||| |  | : :|| ||:: ||  :       :|||| || | :
Db        582 LLQRRAAADSAGKNGLTPLHVAAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQ 641

Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597
               :|  || :||:      : | ||| |  ::| ||:  ||    | :  : |  |
Db        642 MQIASTLLNYGAETNIVTKQGVTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----- 696

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                          | ||||  : :  ||: | :|||| :|  | |  ||  |  ||
Db        697 ----------------TSLHLAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYG 739

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717
              :| :   |:|  | |||  |  :||||:|||:|  | :  :|| ||| |       | | |
Db        740 NVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALA 799

Qy        718 L------VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 771
              :      :|  |   ::|  :  :       |    : :  |    : |          :|
Db        800 IAKRLGYISVVDTLKVVTEEVTTTTTTITEK----HKLNVPETMTEVL--------DVSD 847

Qy        772 ASSLDNLSGSFSEL-----------SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RN 818
                | ::|   |              |: ||   :| : |       |   |:  |   |:
Db        848 EEGDDTMTGDGGEYLRPEDLKELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRS 907
```

```
Qy        819 LGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERL 867
Db        908 HTLSHAS--YLRDSAVMDDSVVIPSHQVSTLAKEAERNSYRLSWGTENL 954
```

RESULT 17
US-09-404-108-4
; Sequence 4, Application US/09404108
; Patent No. 6590077
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/404,108
; CURRENT FILING DATE: 1999-09-23
; EARLIER APPLICATION NUMBER: 09/172,977
; EARLIER FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
US-09-404-108-4

Query Match 14.9%; Score 860.5; DB 4; Length 1839;
Best Local Similarity 29.8%; Pred. No. 2.1e-68;
Matches 283; Conservative 127; Mismatches 386; Indels 153; Gaps 21;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         75 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPL 134

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL--- 115
Db        135 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLP 194

Qy        116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----------TPLHLAAGYNRVKIVQL 163
Db        195 ALHIAARKDDTKSAALLLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 251

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
Db        252 LLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 311

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 283
Db        312 DQVVELLLERGA-PLLARTKN--------------------GLSPLHMAAQGDHVECVK 349

Qy        284 HLSLEMVNFKHP-----QTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPL 338
Db        350 HL----LQHKAPVDDVTLDYLTALHVAA---HCGHYRVTKLLLDKRANPNARALNGFTPL 402

Qy        339 HVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISL 398
Db        403 HIACKKNRIKVMELLVKYGASIQAITESGLTPIHVAAFMGHLNIVLLLLQNGASPDVTNI 462

Qy        399 QGFTALQMGN-----ENVQQLLQEGI-----------------SLGNSEADRQLLE---- 432
Db        463 RGETALHMAARAGQVEVVRCLLRNGALVDARAREEQTPLHIASRLGKTEIVQLLLQHMAH 522

Qy        433 --------------AAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEY 478
Db        523 PDAATTNGYTPLHISAREGQVDVASVLLEAGAAHSLATK-KGFTPLHVAAKYGSLDVAKL 581

Qy        479 LLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGK 538
```

```
Db        582 LLQRRAAADSAGKNGLTPLHVAAHYDNQKVALLLLEKGASPHATAKNGYTPLHIAAKKNQ 641

Qy        539 YEICKLLLQHGADPTKKNRDGNTPLDLV-KDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597
Db        642 MQIASTLLNYGAETNIVTKQGVTPLHLASQEGHTDMVTLLLDKGANIHMSTKSGL----- 696

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
Db        697 ----------------TSLHLAAQEDKVNVADILTKHGADQDAHTKLGYTPLIVACHYG 739

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717
Db        740 NVKMVNFLLKQGANVNAKTKNGYTPLHQAAQQGHTHIINVLLQHGAKPNATTANGNTALA 799

Qy        718 L------VSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSA 771
Db        800 IAKRLGYISVVDTLKVVTEEVTTTTTTITEK----HKLNVPETMTEVL--------DVSD 847

Qy        772 ASSLDNLSGSFSEL-----------SSVVSSSGTEGASSLEKKEVPGVDFSITQFV--RN 818
Db        848 EEGDDTMTGDGGEYLRPEDLKELGDDSLPSSQFLDGMNYLRYSLEGGRSDSLRSFSSDRS 907

Qy        819 LGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERL 867
Db        908 HTLSHAS--YLRDSAVMDDSVVIPSHQVSTLAKEAERNSYRLSWGTENL 954
```

RESULT 18
US-09-172-977-3
; Sequence 3, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: g1841966
US-09-172-977-3

Query Match 14.6%; Score 841.5; DB 2; Length 843;
Best Local Similarity 29.7%; Pred. No. 2.9e-67;
Matches 265; Conservative 138; Mismatches 351; Indels 137; Gaps 21;

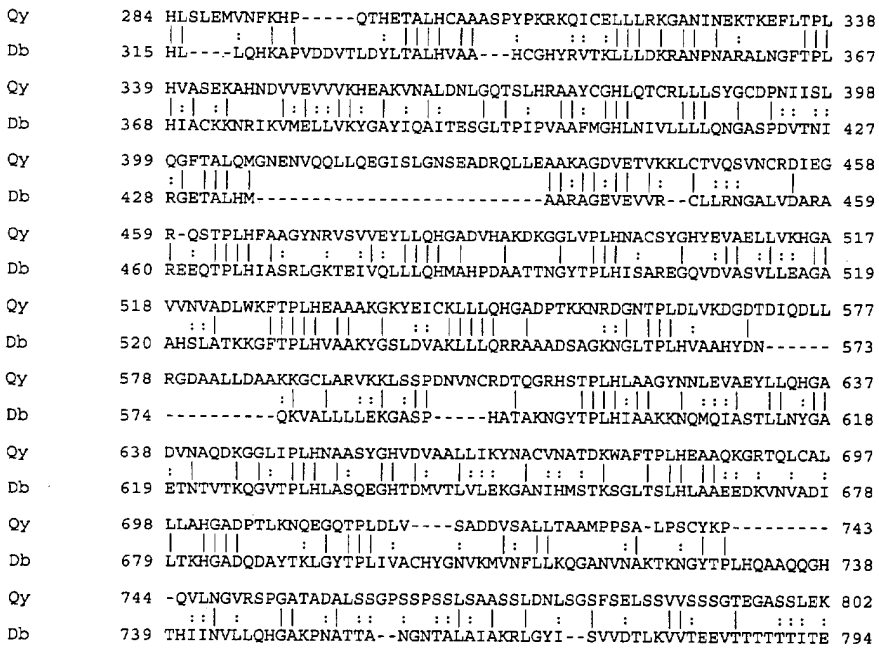
```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         40 GHVGLVQELLGRGSSVDSATKKGNTALHIASLAGQAEVVKVLVKEGANINAQSQNGFTPL 99

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRT----ALDLADPSAKAVLTGEYKKDEL--- 115
Db        100 YMAAQENHIDVVKYLLENGANQSTATEDGFTPLAVALQQGHNQAVAILLENDTKGKVRLP 159

Qy        116 -LESARSGNEEKMMALLTPLNVNCHASDGRKS-----------TPLHLAAGYNRVKIVQL 163
Db        160 ALHIAARKDDTKSAALLLQ---NDHNADVQSKMMVNRTTESGFTPLHIAAHYGNVNVATL 216

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
Db        217 LLNRGAAVDFTARNGITPLHVASKRGNTNMVKLLLDRGGQIDAKTRDGLTPLHCAARSGH 276

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 283
Db        277 DQVVELLLERGA-PLLARTKN--------------------GLSPLHMAAQGDHVECVK 314
```



RESULT 19
US-09-404-108-3
; Sequence 3, Application US/09404108
; Patent No. 6590077
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/404,108
; CURRENT FILING DATE: 1999-09-23
; EARLIER APPLICATION NUMBER: 09/172,977
; EARLIER FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: g1841966
US-09-404-108-3

Query Match 14.6%; Score 841.5; DB 4; Length 843;
Best Local Similarity 29.7%; Pred. No. 2.9e-67;
Matches 265; Conservative 138; Mismatches 351; Indels 137; Gaps 21;





RESULT 20
US-09-031-485-33
; Sequence 33, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation

```
;       STREET:  1825 Sharp Point Drive
;       CITY:  Fort Collins
;       STATE:  Colorado
;       COUNTRY:  USA
;       ZIP:  80525
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect for Windows, Version 7.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/031,485
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/847,429
;       FILING DATE:  24-APR-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Verser, Carol Talkington
;       REGISTRATION NUMBER:  37,459
;       REFERENCE/DOCKET NUMBER:  HW-5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  970/493-7272
;       TELEFAX:  970/484-9505
;   INFORMATION FOR SEQ ID NO:  33:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1745 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-031-485-33

  Query Match             14.2%;  Score 821;  DB 2;  Length 1745;
  Best Local Similarity   29.1%;  Pred. No. 7.8e-65;
  Matches  276;  Conservative  125;  Mismatches  321;  Indels  228;  Gaps   28;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  | |  ||::|  :   :  ||  || |   || |||  ||:  || :|      | |
Db         48 GNLDRVLELLRSGTDINTCNANGLNALHLASKEGHHEVVRELLKRKADVDAATRKGNTAL 107

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP-------------SAKAVLTGE 109
              | |:: |:  :  ||:::||   ::: :| | | :|              :|   |: |
Db        108 HIASLAGQELIVTVLVENGANVNVQSLNGFTPLYMAAQENHESVVRYLLAHNANQALSTE 167

Qy        110 YKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQ-- 166
                |  : : |::  :  ||       : : |:   | ||:||  :  |   ||||
Db        168 DGFTPLAVALQQGHDRVVAVLLE------NDTRGKVRLPALHIAAKKDDTKAATLLLQNE 221

Qy        167 HGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV 226
              | :||    |   ||| |  ||:  | :||:: || ||        :||| |    | :
Db        222 HNSDV--TSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTNM 279

Qy        227 CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLS 286
               ||||::||   :::|  :  :
Db        280 VSLLLAHGA---VIDCRTRDLL----------------------------------- 298

Qy        287 LEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAH 346
                    | ||||: | :    |: :||| ||| |: |||  | |||:|::
Db        299 ------------TPLHCASRSGH---DQVVDLLLEKGAPISAKTKNGLAPLHMAAQ--- 339

Qy        347 NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM 406
              | :| | :            | || ||:|||::  :|||   |||  :| ||| | :
Db        340 --VDDVTVDY----------LTPLHVAAHCGHVRVAKLLLDRNADPNARALNGFTPLHI 386

Qy        407 G--------------NENVQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKK 444
                           :  ::   : |:|       :|       ||:     || ||
Db        387 ACKKNRIKIVELLLKYHAAIEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATV-- 444

Qy        445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 504
                        |  |||| ||  |:  :|  |:::|| | |  :    ||| |   |
Db        445 -------------RGETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIASRLG 490

Qy        505 HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLD 564
```

```
              : ::  ||::  |  | |    :|||| || :|: |:  :|: || | |   : | |||
Db        491 NTDIVILLLQANASPNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLH 550

Qy        565 LVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL---SSPDNVNCRDTQGRHS-TPLHLA 620
              |                      ||| | |   | |    :|       | :|::  ||||:|
Db        551 L--------------------AAKYGNLPVAKSLLERGTP-----VDIEGKNQVTPLHVA 585

Qy        621 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 680
              | ||| :||  ||::||  :|  | |  ||| ||    :|:|: |: | | ||  | |
Db        586 AHYNNDKVALLLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGF 645

Qy        681 TPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL------TAAMPP 734
              |||| |||:|  :: |||: :||    : : | ||: | : :|  ::        ||: |
Db        646 TPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSVAEELVKENAAIDP 705

Qy        735 SALP-------SCYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSE 784
                        :|:  |: | ||     ||     ::    :|   :|
Db        706 KTKAGYTPLHVACHFGQI-NMVRFLIEHGARVSVITRASYTPLHQAAQQG---------- 754

Qy        785 LSSVVSSSGTEGASSLEKKEVPGVDFSITQ----FVRNLG----LEHLMDIFEREQIT-L 835
               :|||      |||        | |  | |         ||    :| |  | |  ||
Db        755 HNSVVRYLLEHGAS-------PNVHTSTGQTPLSIAERLGYVSVVEALKTITETTVITET 807

Qy        836 DVLVEMGHK------------------ELKEIGINAYGH--RHKLIKGV 864
                : |  :|                  |  :|  ||: |     | ||:
Db        808 TTVTEERYKPQNPEAMNETMFSDSEDEGEDNQITANAHAHDFSESLTKGL 857


RESULT 21
US-08-847-429A-33
; Sequence 33, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1745 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-429A-33

  Query Match             14.2%;  Score 821;  DB 2;  Length 1745;
  Best Local Similarity   29.1%;  Pred. No. 7.8e-65;
```

Matches 276; Conservative 125; Mismatches 321; Indels 228; Gaps 28;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         48 GNLDRVLELLRSGTDINTCNANGLNALHLASKEGHHEVVRELLKRKADVDAATRKGNTAL 107

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP------------SAKAVLTGE 109
Db        108 HIASLAGQELIVTVLVENGANVNVQSLNGFTPLYMAAQENHESVVRYLLAHNANQALSTE 167

Qy        110 YKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQ-- 166
Db        168 DGFTPLAVALQQGHDRVVAVLLE------NDTRGKVRLPALHIAAKKDDTKAATLLLQNE 221

Qy        167 HGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV 226
Db        222 HNSDV--TSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTNM 279

Qy        227 CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLS 286
Db        280 VSLLLAHGA---VIDCRTRDLL------------------------------------- 298

Qy        287 LEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAH 346
Db        299 ------------TPLHCASRSGH---DQVVDLLLEKGAPISAKTKNGLAPLHMAAQ--- 339

Qy        347 NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM 406
Db        340 --VDDVTVDY-----------LTPLHVAAHCGHVRVAKLLLDRNADPNARALNGFTPLHI 386

Qy        407 G---------------NENVQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKK 444
Db        387 ACKKNRIKIVELLLKYHAAIEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATV-- 444

Qy        445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 504
Db        445 -------------RGETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIASRLG 490

Qy        505 HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLD 564
Db        491 NTDIVILLLQANASPNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLH 550

Qy        565 LVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL---SSPDNVNCRDTQGRHS-TPLHLA 620
Db        551 L--------------------AAKYGNLPVAKSLLERGTP-----VDIEGKNQVTPLHVA 585

Qy        621 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 680
Db        586 AHYNNDKVALLLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGF 645

Qy        681 TPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL------TAAMPP 734
Db        646 TPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSVAEELVKENAAIDP 705

Qy        735 SALP-------SCYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSE 784
Db        706 KTKAGYTPLHVACHFGQI-NMVRFLIEHGARVSVITRASYTPLHQAAQQG---------- 754

Qy        785 LSSVVSSSGTEGASSLEKKEVPGVDFSITQ----FVRNLG----LEHLMDIFEREQIT-L 835
Db        755 HNSVVRYLLEHGAS-------PNVHTSTGQTPLSIAERLGYVSVVEALKTITETTVITET 807

Qy        836 DVLVEMGHK-----------------ELKEIGINAYGH--RHKLIKGV 864
Db        808 TTVTEERYKPQNPEAMNETMFSDSEDEGEDNQITANAHAHDFSESLTKGL 857
```

RESULT 22
US-09-065-474-33
; Sequence 33, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-33

Query Match 14.2%; Score 821; DB 3; Length 1745;
Best Local Similarity 29.1%; Pred. No. 7.8e-65;
Matches 276; Conservative 125; Mismatches 321; Indels 228; Gaps 28;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         48 GNLDRVLELLRSGTDINTCNANGLNALHLASKEGHHEVVRELLKRKADVDAATRKGNTAL 107

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP------------SAKAVLTGE 109
Db        108 HIASLAGQELIVTVLVENGANVNVQSLNGFTPLYMAAQENHESVVRYLLAHNANQALSTE 167

Qy        110 YKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQ-- 166
Db        168 DGFTPLAVALQQGHDRVVAVLLE------NDTRGKVRLPALHIAAKKDDTKAATLLLQNE 221

Qy        167 HGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV 226
Db        222 HNSDV--TSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTNM 279

Qy        227 CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLS 286
Db        280 VSLLLAHGA---VIDCRTRDLL------------------------------------- 298

Qy        287 LEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAH 346
Db        299 ------------TPLHCASRSGH---DQVVDLLLEKGAPISAKTKNGLAPLHMAAQ--- 339

Qy        347 NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM 406
Db        340 --VDDVTVDY-----------LTPLHVAAHCGHVRVAKLLLDRNADPNARALNGFTPLHI 386

Qy        407 G---------------NENVQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKK 444
Db        387 ACKKNRIKIVELLLKYHAAIEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATV-- 444
```

```
Qy        445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 504
Db        445 --------------RGETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIASRLG 490

Qy        505 HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLD 564
Db        491 NTDIVILLLQANASPNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLH 550

Qy        565 LVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL---SSPDNVNCRDTQGRHS-TPLHLA 620
Db        551 L-------------------AAKYGNLPVAKSLLERGTP-----VDIEGKNQVTPLHVA 585

Qy        621 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 680
Db        586 AHYNNDKVALLLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGF 645

Qy        681 TPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL------TAAMPP 734
Db        646 TPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSVAEELVKENAAIDP 705

Qy        735 SALP-------SCYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSE 784
Db        706 KTKAGYTPLHVACHFGQI-NMVRFLIEHGARVSVITRASYTPLHQAAQQG---------- 754

Qy        785 LSSVVSSSGTEGASSLEKKEVPGVDFSITQ----FVRNLG----LEHLMDIFEREQIT-L 835
Db        755 HNSVVRYLLEHGAS-------PNVHTSTGQTPLSIAERLGYVSVVEALKTITETTVITET 807

Qy        836 DVLVEMGHK-------------------ELKEIGINAYGH--RHKLIKGV 864
Db        808 TTVTEERYKPQNPEAMNETMFSDSEDEGEDNQITANAHAHDFSESLTKGL 857
```

```
RESULT 23
US-09-557-034-33
; Sequence 33, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;  INFORMATION FOR SEQ ID NO: 33:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1745 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-557-034-33

  Query Match           14.2%;  Score 821;  DB 4;  Length 1745;
  Best Local Similarity  29.1%;  Pred. No. 7.8e-65;
  Matches  276;  Conservative  125;  Mismatches  321;  Indels  228;  Gaps   28;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
Db         48 GNLDRVLELLRSGTDINTCNANGLNALHLASKEGHHEVVRELLKRKADVDAATRKGNTAL 107

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADP-------------SAKAVLTGE 109
Db        108 HIASLAGQELIVTVLVENGANVNVQSLNGFTPLYMAAQENHESVVRYLLAHNANQALSTE 167

Qy        110 YKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTP-LHLAAGYNRVKIVQLLLQ-- 166
Db        168 DGFTPLAVALQQGHDRVVAVLLE------NDTRGKVRLPALHIAAKKDDTKAATLLLQNE 221

Qy        167 HGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV 226
Db        222 HNSDV--TSKSGFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNISPLHVATKWGRTNM 279

Qy        227 CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLS 286
Db        280 VSLLLAHGA---VIDCRTRDLL------------------------------------- 298

Qy        287 LEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAH 346
Db        299 -------------TPLHCASRSGH---DQVVDLLLEKGAPISAKTKNGLAPLHMAAQ--- 339

Qy        347 NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQM 406
Db        340 --VDDVTVDY-----------LTPLHVAAHCGHVRVAKLLLDRNADPNARALNGFTPLHI 386

Qy        407 G---------------NENVQQLLQEGIS-------LGNSEADRQLLEAAKAGDVETVKK 444
Db        387 ACKKNRIKIVELLLKYHAAIEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATV-- 444

Qy        445 LCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG 504
Db        445 --------------RGETPLHLAARANQTDIVRVLVRNGAQVDAAARELQTPLHIASRLG 490

Qy        505 HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLD 564
Db        491 NTDIVILLLQANASPNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLH 550

Qy        565 LVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL---SSPDNVNCRDTQGRHS-TPLHLA 620
Db        551 L-------------------AAKYGNLPVAKSLLERGTP-----VDIEGKNQVTPLHVA 585

Qy        621 AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAF 680
Db        586 AHYNNDKVALLLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGF 645

Qy        681 TPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALL------TAAMPP 734
Db        646 TPLHLAAQEGHREMAALLIENGAKVGAQARNGLTPMHLCAQEDRVSVAEELVKENAAIDP 705

Qy        735 SALP-------SCYKPQVLNGVR---SPGATADALSSGPSSPSSLSAASSLDNLSGSFSE 784
Db        706 KTKAGYTPLHVACHFGQI-NMVRFLIEHGARVSVITRASYTPLHQAAQQG---------- 754

Qy        785 LSSVVSSSGTEGASSLEKKEVPGVDFSITQ----FVRNLG----LEHLMDIFEREQIT-L 835
Db        755 HNSVVRYLLEHGAS-------PNVHTSTGQTPLSIAERLGYVSVVEALKTITETTVITET 807

Qy        836 DVLVEMGHK-------------------ELKEIGINAYGH--RHKLIKGV 864
```

```
Db        808 TTVTEERYKPQNPEAMNETMFSDSEDEGEDNQITANAHAHDFSESLTKGL 857


RESULT 24
US-10-164-595-38
; Sequence 38, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
;  APPLICANT: OriGene Technologies, Inc
;  TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
;  FILE REFERENCE: 1U 103 R1
;  CURRENT APPLICATION NUMBER: US/10/164,595
;  CURRENT FILING DATE:  2002-06-10
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
;   LENGTH: 994
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-164-595-38

  Query Match             11.8%;  Score 681;  DB 4;  Length 994;
  Best Local Similarity   28.3%;  Pred. No. 1.7e-52;
  Matches  239;  Conservative  114;  Mismatches  310;  Indels  182;  Gaps   26;

Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62
              |  : :  |:     |   |      ||| |   | ||:: ||:  ||  ||:||   |||
Db         19 GDPEEIRMLIHKTEDVNTLDSEKRTPLHVAAFLGDAEIIELLILSGARVNAKDNMWLTPL 78

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSG 122
              | |     :   ||::| |:   |: : :| | :|  : |||
Db         79 HRAVASRSEEAVQVLIKHSADVNARDKNWQTPLHVA-AANKAV---------------- 120

Qy        123 NEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPL 182
                 |   :: ||  : : ||    | || ||    |::| ||| ||:::| || |   |
Db        121 ---KCAEVIIPLLSSVNVSDRGGRTALHHAALNGHVEMVNLLLAKGANINAFDKKDRRAL 177

Qy        183 HNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNC 242
              | |   || :|  ||: ||| |    |   :|||| |||  :: |   ||: | :   :|
Db        178 HWAAYMGHLDVVALLINHGAEVTCKDKKGYTPLHAAASNGQINVVKHLLNLGVEIDEINV 237

Qy        243 HNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHE-TA 301
              :  :|: :|           |           ::| |  :  :      |   | :  |
Db        238 YGNTALHIA---------CYN---------GQDAVVNELIDY----GANVNQPNNNGFTP 275

Qy        302 LHCAAASPYPKRKQIC-ELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV 360
              || |||| :     :| |||:  ||::| ::|:  :|||: :         : ::::  ::
Db        276 LHFAAASTH---GALCLELLVNNGADVNIQSKDGKSPLHMTAVHGRFTRSQTLIQNGGEI 332

Qy        361 NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG-----FTALQMGNENVQQLL 415
              : :|  | | || ||  ||      |:: | |    :          ||  ::  ::||
Db        333 DCVDKDGNTPLHVAARYGHELLINTLITSGADTAKCGIHSMFPLHLAALNAHSDCCRKLL 392

Qy        416 QEG-----ISLGNSEADRQLLEAAKAG-DVETVKKLCTVQSVNCRDIEGRQSTPLHFAAG 469
                  :|| ::|           || :::|  |              ||  | || ||
Db        393 SSGQKYSIVSLFSNE------HVLSAGFEIDTPDKF-----------GR--TCLHAAAA 432

Qy        470 YNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTP 529
                |  :: |   ||| | ||| |  ||| | :  |:   | ||  || ||  | |  |
Db        433 GGNVECIKLLQSSGADFHKKDKCGRTPLHYAAANCHFHCIETLVTTGANVNETDDWGRTA 492

Qy        530 LHEAAA--------------------------KGKYEIC-KLLLQHGADPTKKNRDG-- 559
              || |||                          : :  :| : |||: |:|: ::::|
Db        493 LHYAAASDMDRNKTILGNAHDNSEELERARELKEKEATLCLEFLLQNDANPSIRDKEGYN 552

Qy        560 --------------------------------NTPLDLVK-DG-------------DTD 572
                                             :|| |   :|              | |
Db        553 SIHYAAAYGHRQCLELLLERTNSGFEESDSGATKSPLHLAAYNGHHQALEVLLQSLVDLD 612

Qy        573 IQDLLRGDAALLDAAKKGCLARVKKLSS-------PDNVNCRDTQGRHSTPLHLAAGYNN 625
              |:|  :|  ||  || ||   |: | :         |||  |       ||||  |:  |
Db        613 IRD-EKGRTALDLAAFKGHTECVEALINQGASIFVKDNVTKR-------TPLH-ASVING 663


Qy        626 LEVAEYLLQHGAD----VNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFT 681
                :   ||   ||    |: :|  |  ||  | :|||:|  :||::  | |:  |   |
Db        664 HTLCLRLLLEIADNPEAVDVKDAKGQTPLMLAVAYGHIDAVSLLLEKEANVDTVDILGCT 723

Qy        682 PLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADD----VSALLTAAMPPSAL 737
              ||      |  :   :||          |:  |:|||   :|       :| ||  |:  |
Db        724 ALHRGIMTGHEECVQMLLEQEVSILCKDSRGRTPLHYAAARGHATWLSELLQMAL--SEE 781

Qy        738 PSCYK 742
                |:|
Db        782 DCCFK 786


RESULT 25
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
;  APPLICANT: Morrow, Jon S.
;  APPLICANT: Devarajan, Prasad
;  TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identificat
;  FILE REFERENCE: 44574-5002-US
;  CURRENT APPLICATION NUMBER: US/09/082,059A
;  CURRENT FILING DATE: 1998-05-21
;  EARLIER APPLICATION NUMBER: 60/047356
;  EARLIER FILING DATE: 1997-05-21
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1088
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-082-059-2

  Query Match             10.9%;  Score 629.5;  DB 3;  Length 1088;
  Best Local Similarity   28.1%;  Pred. No. 9.6e-48;
  Matches  195;  Conservative   85;  Mismatches  234;  Indels  181;  Gaps   16;

Qy        164 LLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR 223
              ||    |: :||       ||| ||      :| |||:|||| : |:     ||:| ||
Db          7 LLDKKANPNAKALNGFTPLHIACKKNRIKVMELLLKHGASIQAVTESGLTPIHVAAFMGH 66

Qy        224 VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKK 283
              | : | |: :|| |   |
Db         67 VNIVSQLMHHGASPNTTNVRG---------------------------------------- 87

Qy        284 HLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASE 343
                                  |||| || |    : ::   |:: || :  | |:  |||:::
Db         88 ---------------ETALHMAARS---GQAEVVRYLVQDGAQVEAKAKDDQTPLHISAR 129

Qy        344 KAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTA 403
                |:|: :::  |  ||    | | || :|  ||     || :|   :| : :||
Db        130 LGKADIVQQLLQQGASPNAATTSGYTPLHLSAREGHEDVAAFLLDHGASLSITTKKGF-- 187

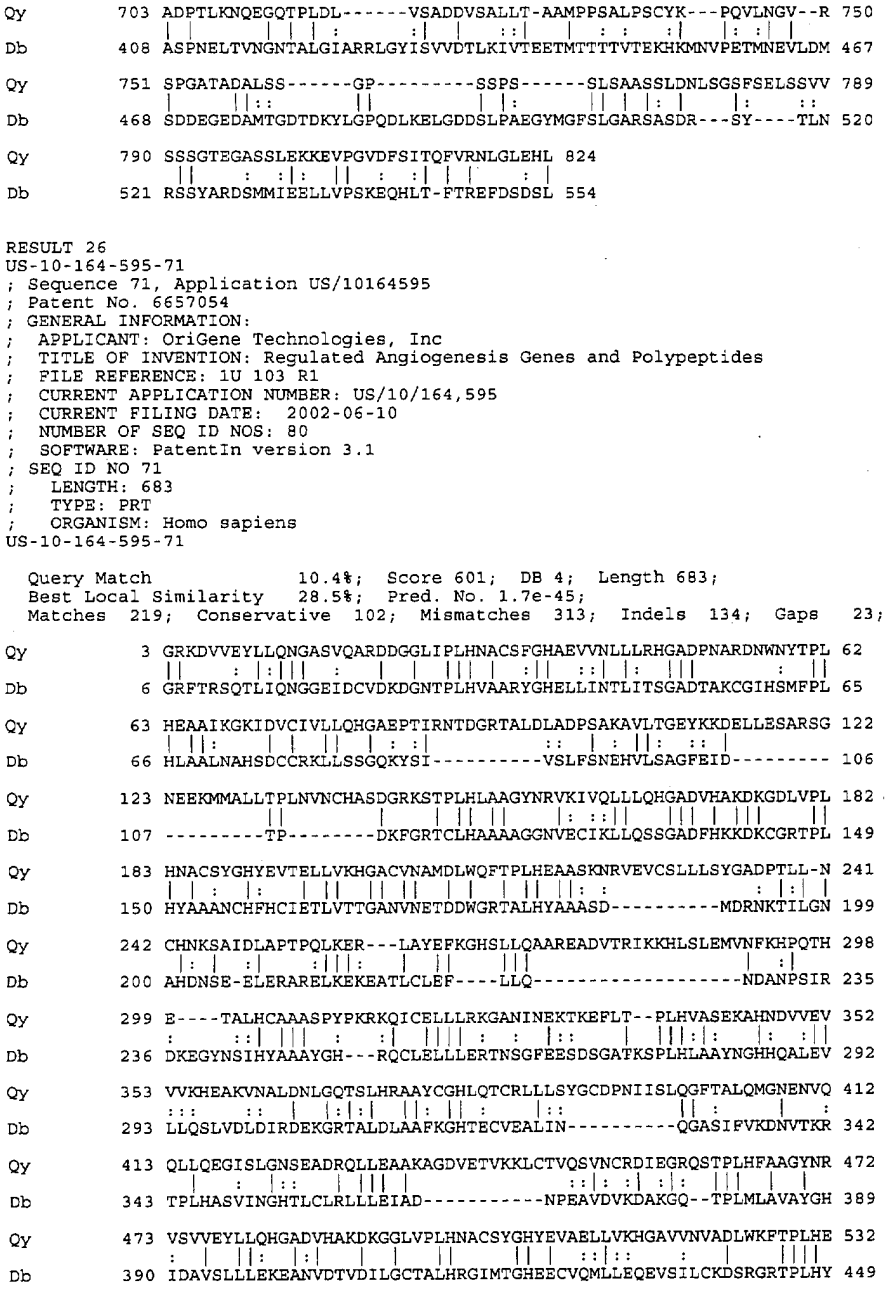
Qy        404 LQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTP 463
                                                                    ||
Db        188 ------------------------------------------------------TP 189

Qy        464 LHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVAD 523
              || ||  | :: |   |||  |   | || ||| |  | : :|| ||:  ||  : |
Db        190 LHVAAKYGKLEVANLLLQKSASPDAAGKSGLTPLHVAAHYDNQKVALLLLDQGASPHAAA 249

Qy        524 LWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAA 582
                :|||| || | : :|   ||::|||      | |   : |  ::|  |:  || | |
Db        250 KNGYTPLHIAAKKNQMDIATTLLEYGADANAVTRQGIASVHLAAQEGHVDMVSLLLGRNA 309

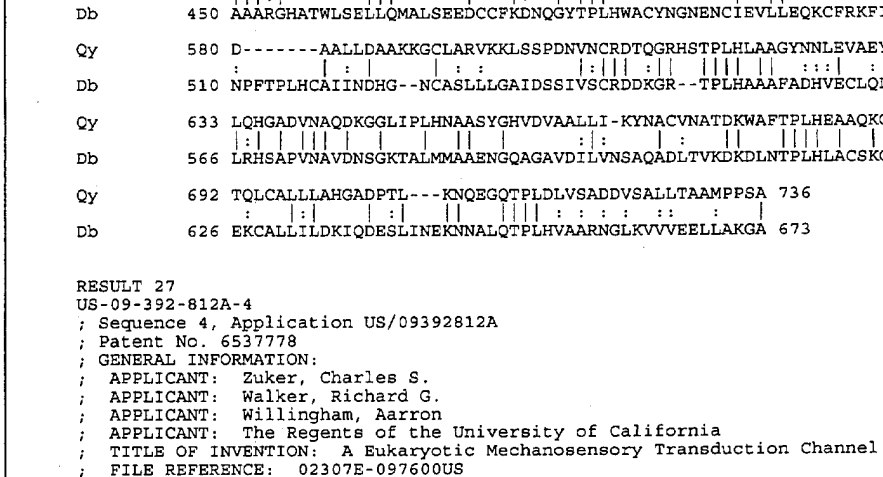
Qy        583 LLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQ 642
                                  |||  :  |  |||||||  : : ||| |:  || |:||
Db        310 --------------------NVNLSNKSG--LTPLHLAAQEDRVNVAEVLVNQGAHVDAQ 347

Qy        643 DKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHG 702
              | |  |||    ||:: :   |::::| |||   |  :||||:|||:|  | :  :|| :
Db        348 TKMGYTPLHVGCHYGNIKIVNFLLQHSAKVNAKTKNGYTPLHQAAQQGHTHIINVLLQNN 407
```



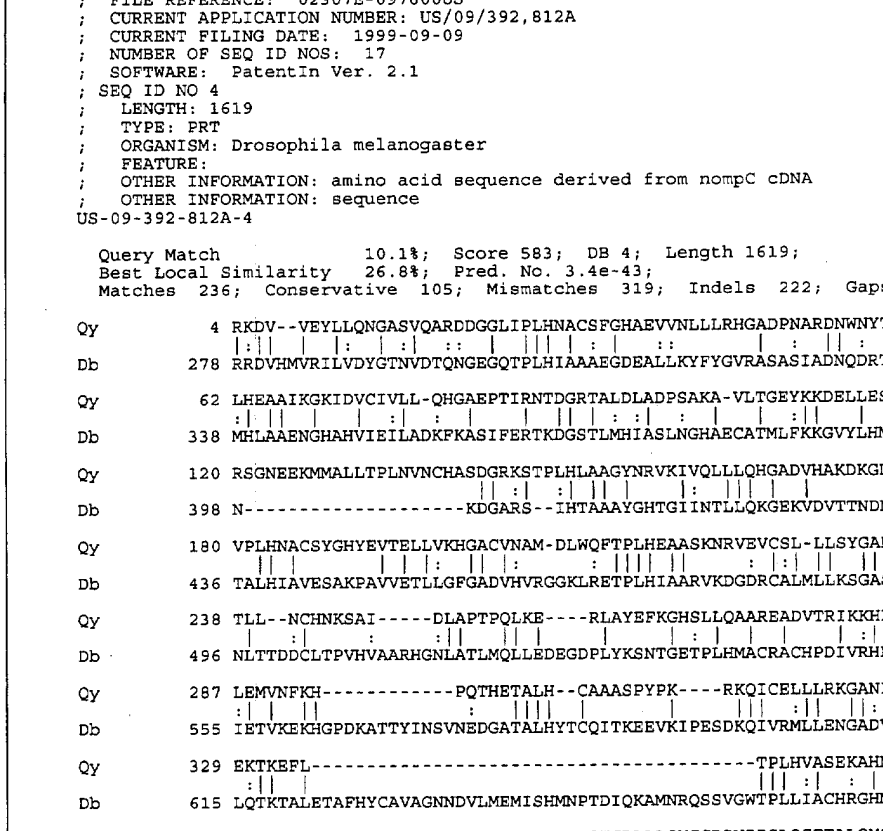
RESULT 26
US-10-164-595-71
; Sequence 71, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-595-71

Query Match 10.4%; Score 601; DB 4; Length 683;
Best Local Similarity 28.5%; Pred. No. 1.7e-45;
Matches 219; Conservative 102; Mismatches 313; Indels 134; Gaps 23;



RESULT 27
US-09-392-812A-4
; Sequence 4, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/09/392,812A
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1619
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC cDNA
; OTHER INFORMATION: sequence
US-09-392-812A-4

Query Match 10.1%; Score 583; DB 4; Length 1619;
Best Local Similarity 26.8%; Pred. No. 3.4e-43;
Matches 236; Conservative 105; Mismatches 319; Indels 222; Gaps 26;



```
Db        675 LVNNLLANHARVDVFDTEGRSALHLAAERGYLHVCDALLTNKAFINSKSRVGRTALHL-- 732

Qy        409 ENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAA 468

Db        733 ------------------------AAMNGFTHLVKFLIKDHNAVIDILTLRKQTPLHLAA 768

Qy        469 GYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL-LVKHGAVVN------- 520

Db        769 ASGQMEVCQLLLELGANIDATDDLGQKPIHVAAQNNYSEVAKLFLQQHPSLVNATSKDGN 828

Qy        521 --------------VADLWKF---------------TPLHEAAAKGKYEICKLLLQHGAD 551

Db        829 TCAHIAAMQGSVKVIEELMKFDRSGVISARNKLTDATPLQLAAEGGHADVVKALVRAGAS 888

Qy        552 PTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKK------------GCLARVKKL 598

Db        889 CTEENKAGFTAVHLAAQNGHGQVLDVLKSTNSLRINSKKLGLTPLHVAAYYGQADTVREL 948

Qy        599 --SSPDNVNCRDTQGRH----------STPLHLAA------------------------ 621

Db        949 LTSVPATVKSETPTGQSLFGDLGTESGMTPLHLAAFSGNENVVRLLLNSAGVQVDAATIE 1008

Qy        622 -GYNNLEVAEY---------LLQHGAD-VNAQDKGGLIPLHNAASYGHVDVAALLIKYNA 670

Db       1009 NGYNPLHLACFGGHMSVVGLLLSRSAELLQSQDRNGRTGLHIAAMHGHIQMVEILLGQGA 1068

Qy        671 CVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEG 712

Db       1069 EINATDRNGWTPLHCAAKAGHLEVVKLLCEAGASPKSETNYG 1110
```

RESULT 28
US-09-392-812A-2
; Sequence 2, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/09/392,812A
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1704
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-09-392-812A-2

Query Match 9.4%; Score 541; DB 4; Length 1704;
Best Local Similarity 24.8%; Pred. No. 2.5e-39;
Matches 201; Conservative 115; Mismatches 273; Indels 220; Gaps 22;

```
Qy          3 GRKDVVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDNWNYTPL 62

Db         45 GHAECATMLFKKGVYLHMPNKDGARSIHTAAAYGHTGIINTLLQKGEKVDVTTN-NYTAL 103

Qy         63 HEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESAR-- 120

Db        104 HIAVESAKPAVVETLLGFGADVHVRG------------------GKLRETPLHIAARVK 144

Qy        121 SGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGY-NRVKIVQLLLQHGADVHAKDKGDL 179

Db        145 DGDRCALMLLKSGASPNLTTDD--CLTPVHVAARHGNLATLMQLLEDEGDPLYKSNTGE- 201

Qy        180 VPLHNACSYGHYEVTELLV-----KHG-----ACVNAMDLWQFTPLHEAASKNRVEV--- 226

Db        202 TPLHMACRACHPDIVRHLIETVKEKHGPDKATTYINSVNEDGATALHYTCQITKEEVKIP 261

Qy        227 ------CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTR 280

Db        262 ESDKQIVRMLLENGADVTL---QTKTALETA---------------FHYCAVAGNNDVL- 302

Qy        281 IKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHV 340

Db        303 ------MEMISHMNPTDIQKAM------------------------NRQSSVGWTPLLI 331

Qy        341 ASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQG 400

Db        332 ACHRGHMELVNNLLANHARVDVFDTEGRSALHLAAERGYLHVCDALLTNKAFINSKSRVG 391

Qy        401 FTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQ 460

Db        392 RTALHL------------------------AAMNGFTHLVKFLIKDHNAVIDILTLRK 425

Qy        461 STPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAEL-LVKHGAVV 519

Db        426 QTPLHLAAASGQMEVCQLLLELGANIDATDDLGQKPIHVAAQNNYSEVAKLFLQQHPSLV 485

Qy        520 N---------------------VADLWKF---------------TPLHEAAAKGKYEICK 543

Db        486 NATSKDGNTCAHIAAMQGSVKVIEELMKFDRSGVISARNKLTDATPLQLAAEGGHADVVK 545

Qy        544 LLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPD 602

Db        546 ALVRAGASCTEENKAGFTAVHLAAQNGHGQVLDVLKSTNSLRINSKKLGL---------- 595

Qy        603 NVNCRDTQGRHSTPLHLAAGYNNLE-VAEYLLQHGADVNAQ------------DKGGLIP 649

Db        596 ------------TPLHVAAYYGQADTVRELLTSVPATVKSETPTGQSLFGDLGTESGMTP 643

Qy        650 LHNAA------------------------SYGHVDVAALLIKYNACVNATDKWAFTPL 683

Db        644 LHLAAFSGNENVVRLLLNSAGVQVDAATIENMHGHIQMVEILLGQGAEINATDRNGWTPL 703

Qy        684 HEAAQKGRTQLCALLLAHGADPTLKNQEG 712

Db        704 HCAAKAGHLEVVKLLCEAGASPKSETNYG 732
```

RESULT 29
US-09-392-812A-6
; Sequence 6, Application US/09392812A
; Patent No. 6537778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Walker, Richard G.
; APPLICANT: Willingham, Aarron
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: A Eukaryotic Mechanosensory Transduction Channel
; FILE REFERENCE: 02307E-097600US
; CURRENT APPLICATION NUMBER: US/09/392,812A
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1709
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: amino acid sequence derived from nompC genomic
; OTHER INFORMATION: sequence
US-09-392-812A-6

Query Match 9.3%; Score 536.5; DB 4; Length 1709;
Best Local Similarity 24.7%; Pred. No. 6.4e-39;
Matches 237; Conservative 108; Mismatches 317; Indels 297; Gaps 32;

```
Qy          7 VVEYLLQNGASVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNAR--DNWNYTPLHE 64
```

```
Db        237 IVKRILEDTDQRMAEDGDGSLPIHLAFKFGNVNIVELLLSGPSDEQTRKADGNGDTLLHL 296

Qy         65 AAIKGKIDVCIVLLQHGAE-PTIRNTDGRTALDLADPSAKAVLTGEYKKDELLESARSGN 123

Db        297 AARSGNIEAVRTAIAAGCDNANVQNRVGRTPLHEC-----LTVTGTQK-----GYVAEVG 346

Qy        124 EEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL-QHGADVHAK-------- 174

Db        347 DQNMLKIMFKLRADANIHDKEDKTPVHVAAERGDTSMVESLIDKFGGSIRARTRDGSTLL 406

Qy        175 ------------------------DKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW 210

Db        407 HIAACSGHTSTALAFLKRVPLFMPNKKGALGLHSAAAAGFNDVVKMLIARGTNVDVRTRD 466

Qy        211 QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL 270

Db        467 NYTALHVAVQSGKASVVETLLGSGAD-------------------------IHVKGGELM 501

Qy        271 QAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEK 330

Db        502 DG--------------------------ETCLHIAARS---GNKDIM-LLLDENADSKIS 531

Qy        331 TKEFLTPLHVASEKAHNDVVEVVVKH----------EAKVNALDNLGQTSLHRAAYCGHL 380

Db        532 SKIGETPLQVAAKSCNFEAASMILKHLSEVLTQEQLKEHVNHRTNDGFTALHYAAEIEQR 591

Qy        381 Q---------TCRLLLSYGCDPNIISLQ-GFTALQM----GNENVQQLLQEGISLG---- 422

Db        592 QLHFPGEDAKLVNLLIDYGGMVEMPSLNANETAMHMAARSGNQAVLLAMVNKIGAGAVQI 651

Qy        423 --NSEADR---QLLEAAKAGD------VETVKKLC------------------------- 446

Db        652 VQNKQSKNGWSPLLEACARGHSGVANILLKVLVLCVGPGPGPGPRLQGRGYWTRTRARVT 711

Qy        447 ----------------------------------------TV--------QSVNCRDIE 457

Db        712 VPWLQYQGYWARTRTRTRARATGPGLQDQGYWARTRTRTKVTVPRLLGDHHARIDVFDEM 771

Qy        458 GRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVK-HG 516

Db        772 GR--TALHLAAFNGHLSLVHLLLQHKAFVNSKSKTGEAPLHLAAQHGHVKVVNVLVQDHG 829

Qy        517 AVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGD-TDIQD 575

Db        830 AALEAITLDNQTALHFAAKFGQLAVSQTLLALGANPNARDDKGQTPLHLAAENDFPDVVK 889

Qy        576 LL-------RGDAALLD---------AAKKGCLARVKKLSSPDN--VNCRDTQGRHSTPL 617

Db        890 LFLKMRNNNRSVLTAIDHNGFTCAHIAAMKGSLAVVRELMMIDKPMVIQAKTKTLEATTL 949

Qy        618 HLAAGYNNLEVAEYLLQHGADVNAQD--------------------------------K 644

Db        950 HMAAAGGHANIVKILLENGA--NAEDENSGMTALHLGAKNGFISILEAFDKILWKRCSRK 1007

Qy        645 GGLIPLHNAASYGHVDVAALLIK-----------YNACVNA--TDKWAFTPLHEAAQKG 690

Db       1008 TGLNALHIAAFYGNSDFVNEMLKHVQATVRSEPPIYNHHVNKEFSTEYGFTPLHLAAHSG 1067

Qy        691 RTQLCALLLAHG--ADPT----LKNQEGQTPLDLVSAD------DVSALLTAAMPPSAL 737

Db       1068 HDSLVRMLLNQGVQVDATSTTMMSEKEKERAKDLLNVAVFSENMAVELLITATEYNAAL 1126


RESULT 30
US-09-065-474-139
; Sequence 139, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  139:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  352 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-09-065-474-139

  Query Match             8.9%;  Score 513;  DB 3;  Length 352;
  Best Local Similarity   30.2%;  Pred. No. 5.4e-38;
  Matches  127;  Conservative   64;  Mismatches  144;  Indels   86;  Gaps    5;

Qy        299 ETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEA 358

Db         12 ETPLHLAARA---NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANA 68

Qy        359 KVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEG 418

Db         69 SPNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLHL------------ 116

Qy        419 ISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVE 477

Db        117 --------------AAKYGNLPVAKSL--LERGTPVDIEGKNQVTPLHVAAHYNNDKVAL 160

Qy        478 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 537

Db        161 LLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGFTPLHLAAQEG 220

Qy        538 KYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597

Db        221 HREMAALLIENGAKVGAQARNG---------------------------------------- 242

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657

Db        243 ---------------LTPMHLCAQEDRVSVAEELVKENAAIDPKTKAGYTPLHVACHFG 286

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717

Db        287 QINMVRFLIEHGARVSVITRASYTPLHQAAQQGHNSVVRYLLEHGASPNVHTSTGQTPLS 346

Qy        718 L 718
               :
Db        347 I 347


RESULT 31
US-09-557-034-139
```

```
; Sequence 139, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 139:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 352 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: Protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 139:
US-09-557-034-139

  Query Match             8.9%;  Score 513;  DB 4;  Length 352;
  Best Local Similarity   30.2%;  Pred. No. 5.4e-38;
  Matches  127;  Conservative   64;  Mismatches  144;  Indels   86;  Gaps    5;

Qy        299 ETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEA 358
              || || || :    :  |  :|:| || ::   :|  ||||:||   : |:| ::::  |
Db         12 ETPLHLAARA---NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANA 68

Qy        359 KVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEG 418
              ||      | || ||   |  :   :|: :| |  ::: :||| | :
Db         69 SPNAATRDLYTPLHIAAKEGQEEVAAILMDHGTDKTLLTKKGFTPLHL------------ 116

Qy        419 ISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR-QSTPLHFAAGYNRVSVVE 477
                        ||| |::   | |  ::     ||||: | |||| || ||   |
Db        117 --------------AAKYGNLPVAKSL--LERGTPVDIEGKNQVTPLHVAAHYNNDKVAL 160

Qy        478 YLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKG 537
              ||::||  ||  | |  ||| |      ::|  |: : |  |          ||||| || :|
Db        161 LLLENGASAHAAAKNGYTPLHIAAKKNQMDIASTLLHYKANANAESKAGFTPLHLAAQEG 220

Qy        538 KYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKK 597
              |:  ||:::||    : |:|
Db        221 HREMAALLIENGAKVGAQARNG---------------------------------- 242

Qy        598 LSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                          ||:|| |  : : ||| |::  | :: : | |  ||| | :|
Db        243 ---------------LTPMHLCAQEDRVSVAEELVKENAAIDPKTKAGYTPLHVACHFG 286

Qy        658 HVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLD 717
              :::   ||:: | |:   : ::||||:|||:|   :   || ||| | :    |||||
Db        287 QINMVRFLIEHGARVSVITRASYTPLHQAAQQGHNSVVRYLLEHGASPNVHTSTGQTPLS 346

Qy        718 L 718
              :
Db        347 I 347


RESULT 32
US-09-972-115A-22
; Sequence 22, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 90
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-115A-22

  Query Match             7.6%;  Score 440;  DB 4;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-32;
  Matches   90;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        729 TAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSV 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSV 60

Qy        789 VSSSGTEGASSLEKKEVPGVDFSITQFVRN 818
              ||||||||||||||||||||||||||||||
Db         61 VSSSGTEGASSLEKKEVPGVDFSITQFVRN 90


RESULT 33
US-09-031-485-23
; Sequence 23, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
```

```
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/847,429
;       FILING DATE:  24-APR-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Verser, Carol Talkington
;       REGISTRATION NUMBER:  37,459
;       REFERENCE/DOCKET NUMBER:  HW-5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  970/493-7272
;       TELEFAX:  970/484-9505
;   INFORMATION FOR SEQ ID NO:  23:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  303 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-031-485-23

  Query Match             7.3%;  Score 421.5;  DB 2;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 9e-30;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        327 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
              ::: | ::||||||                                  ||:|||::  :||
Db          1 VDDVTVDYLTPLHV---------------------------------AAHCGHVRVAKLL 27

Qy        387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
              |    |||  :| ||| | :                              | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI-------------------------ACKKNRIKIVELLL 61

Qy        447 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 503
                :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        504 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 563
              ::  :||::|| |: |     |||| |:  |  :|  |||| | |    || |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        564 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 619
              :  |:|  ::  :|    |  ||   |||                          |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        620 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 679
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        680 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 710
              :|||| ||:|  :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 34
US-08-847-429A-23
; Sequence 23, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  303 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-429A-23

  Query Match             7.3%;  Score 421.5;  DB 2;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 9e-30;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        327 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
              ::: | ::||||||                                  ||:|||::  :||
Db          1 VDDVTVDYLTPLHV---------------------------------AAHCGHVRVAKLL 27

Qy        387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
              |    |||  :| ||| | :                              | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI-------------------------ACKKNRIKIVELLL 61

Qy        447 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 503
                :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        504 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 563
              ::  :||::|| |: |     |||| |:  |  :|  |||| | |    || |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        564 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 619
              :  |:|  ::  :|    |  ||   |||                          |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        620 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 679
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        680 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 710
              :|||| ||:|  :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 35
US-09-065-474-23
; Sequence 23, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
```

```
;       COUNTRY:  USA
;       ZIP:  80525
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect for Windows, Version 7.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/065,474
;       FILING DATE:  24-APR-1998
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Verser, Carol Talkington
;       REGISTRATION NUMBER:  37,459
;       REFERENCE/DOCKET NUMBER:  HW-5-C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  970/493-7272
;       TELEFAX:  970/484-9505
;   INFORMATION FOR SEQ ID NO:  23:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  303 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-065-474-23

  Query Match             7.3%;  Score 421.5;  DB 3;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 9e-30;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        327 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
              ::: | ::|||||                                ||:|||::  :||
Db          1 VDDVTVDYLTPLHV--------------------------------AAHCGHVRVAKLL 27

Qy        387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
              |   |||  :| ||| | :                            | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI------------------------ACKKNRIKIVELLL 61

Qy        447 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 503
                :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        504 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 563
              ::  :||::|| |: |     |||| |:  |  :|  |||   | |    ||  |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        564 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 619
               :  |:|  ::  :|     | ||  |||                      |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        620 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 679
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        680 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 710
              :|||| ||:| :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 36
US-09-557-034-23
; Sequence 23, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 23:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 303 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-557-034-23

  Query Match             7.3%;  Score 421.5;  DB 4;  Length 303;
  Best Local Similarity   30.4%;  Pred. No. 9e-30;
  Matches  119;  Conservative   52;  Mismatches  125;  Indels   95;  Gaps    8;

Qy        327 INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLL 386
              ::: | ::|||||                                ||:|||::  :||
Db          1 VDDVTVDYLTPLHV--------------------------------AAHCGHVRVAKLL 27

Qy        387 LSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC 446
              |   |||  :| ||| | :                            | |   :: |: |
Db         28 LDRNADPNARALNGFTPLHI------------------------ACKKNRIKIVELLL 61

Qy        447 TVQSVNCRDIEGRQS---TPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSY 503
                :     ||       :||| ||    :::| |||| ||:       |  ||| |
Db         62 KYHAA----IEATTESGLSPLHVAAFMGAINIVIYLLQQGANADVATVRGETPLHLAARA 117

Qy        504 GHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPL 563
              ::  :||::|| |: |     |||| |:  |  :|  |||   | |    ||  |||
Db        118 NQTDIVRVLVRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQANASPNAATRDLYTPL 177

Qy        564 DL-VKDGDTDIQDLLRG---DAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHL 619
               :  |:|  ::  :|     | ||  |||                      |||||
Db        178 HIAAKEGQEEVAAILMDHGTDKTLL--TKKG----------------------FTPLHL 212

Qy        620 AAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWA 679
              || | || ||: ||: |  |: : |  : ||| || | :  || ||::  |  :|  |
Db        213 AAKYGNLPVAKSLLERGTPVDIEGKNQVTPLHVAAHYNNDKVALLLLENGASAHAAAKNG 272

Qy        680 FTPLHEAAQKGRTQLCALLLAHGADPTLKNQ 710
              :|||| ||:| :  : : || : |:   :::
Db        273 YTPLHIAAKKNQMDIASTLLHYKANANAESK 303


RESULT 37
US-09-031-485-38
; Sequence 38, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
```

```
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  302 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-031-485-38

  Query Match             7.1%;  Score 409;  DB 2;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 1.2e-28;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        300 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 359
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        360 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 419
              : |     | : || ||: | :      ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------- 106

Qy        420 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 479
                                                 |||| ||  |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        480 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 539
              :::|| | |  :    ||| |   |: ::  ||::  |  | |    :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        540 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
              |:  :|: ||:|  |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        599 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                      | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        658 HVDVAALLIKYNACVNATDK 677
                :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302


RESULT 38
```

```
US-08-847-429A-38
; Sequence 38, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  302 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-429A-38

  Query Match             7.1%;  Score 409;  DB 2;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 1.2e-28;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        300 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 359
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        360 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 419
              : |     | : || ||: | :      ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------- 106

Qy        420 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 479
                                                 |||| ||  |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        480 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 539
              :::|| | |  :    ||| |   |: ::  ||::  |  | |    :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        540 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
              |:  :|: ||:|  |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        599 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                      | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        658 HVDVAALLIKYNACVNATDK 677
                :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302
```

```
RESULT 39
US-09-065-474-38
; Sequence 38, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  302 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-065-474-38

  Query Match             7.1%;  Score 409;  DB 3;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 1.2e-28;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        300 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 359
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        360 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 419
              : |    | : || ||: | :     ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------ 106

Qy        420 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 479
                                                  ||||  ||  |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        480 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 539
              :::|| | |  :    ||| |   |: ::  ||::  |  | |    :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        540 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
              |:  :|: ||:|  |   : | ||| |  | |:  :  ||          ::| |
Db        186 EVAAILIDHGSDKTLLTKKGFTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        599 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657
                    | :|::  ||||:|| ||| :||  ||: ||  :|  | |  ||| ||
Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        658 HVDVAALLIKYNACVNATDK 677
                :|:|: |: | |  ||  |
Db        283 QMDIASTLLHYRANANAESK 302


RESULT 40
US-09-557-034-38
; Sequence 38, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;  INFORMATION FOR SEQ ID NO: 38:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 302 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-557-034-38

  Query Match             7.1%;  Score 409;  DB 4;  Length 302;
  Best Local Similarity   29.5%;  Pred. No. 1.2e-28;
  Matches  112;  Conservative   56;  Mismatches  124;  Indels   88;  Gaps    6;

Qy        300 TALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAK 359
              | || ||   : :   : :||| : |: | :     ||||:| :|    |||:::|: |
Db          9 TPLHVAAHCGHVR---VAKLLLDRNADSNARALNGFTPLHIACKKNRIKVVELLLKYHAA 65

Qy        360 VNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGI 419
              : |    | : || ||: | :     ||  | : |: :::|
Db         66 IEATTESGLSPLHVAAFMGAINIVIYLLQQGANANVATVRG------------------ 106

Qy        420 SLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYL 479
                                                  ||||  ||  |:  :|  |
Db        107 ----------------------------------------ETPLHLAARANQTDIVRVL 125

Qy        480 LQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKY 539
              :::|| | |  :    ||| |   |: ::  ||::  |  | |    :| || || :|:
Db        126 VRNGAQVDAAARELQTPLHIASRLGNTDIVILLLQADASPNAATRDLYTLLHIAAKEGQE 185

Qy        540 EICKLLLQHGADPTKKNRDGNTPLDL-VKDGDTDIQDLLRGDAALLDAAKKGCLARVKKL 598
```

```
Db        186 EVAAILIDHGSDKTLLTKKGPTPLHLAAKYGNLPVAKLL---------LERGTLV----- 231

Qy        599 SSPDNVNCRDTQGRHS-TPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYG 657

Db        232 ---------DIEGKNQVTPLHVAAHYNNDKVALLLLESGASAHAVAKNGYTPLHIAAKKN 282

Qy        658 HVDVAALLIKYNACVNATDK 677

Db        283 QMDIASTLLHYRANANAESK 302


RESULT 41
US-09-031-485-28
; Sequence 28, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-031-485-28

  Query Match             6.8%;  Score 392;  DB 2;  Length 348;
  Best Local Similarity   26.5%;  Pred. No. 5.6e-27;
  Matches  116;  Conservative   56;  Mismatches  142;  Indels  124;  Gaps    6;

Qy        291 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 350

Db         28 NSQHSNKGESSASFLRAARAGNLDRVLE-LLRSGTDINTCNANGLNALHLASKEGHHEVV 86

Qy        351 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 410

Db         87 RELLKRKADVDAATRKGNTALHIASLAGQELIVTVLVENGANVNVQSLNGF--------- 137

Qy        411 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 470

Db        138 ----------------------------------------------------TPLYMAAQE 146
```

```
Qy        471 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 530

Db        147 NHESVVRYLLAHNANQALSTEDGFTPLAVALQQGHDRVVAVLLENDTRGKV----RLPAL 202

Qy        531 HEAAAKGKYEICKLLLQ--HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 588

Db        203 HIAAKKDDTKAATLLLQNEHNSDVTSKS------------------------------- 230

Qy        589 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 648

Db        231 -----------------------GFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNIS 266

Qy        649 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 708

Db        267 PLHVATKWGRTNMVSLLLAHGAVIDCRTRDLLTPLHCASRSGHDQVVDLLLEKGAPISAK 326

Qy        709 NQEGQTPLDLVS-ADDVS 725

Db        327 TKNGLAPLHMAAQVDDVT 344


RESULT 42
US-08-847-429A-28
; Sequence 28, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  348 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-847-429A-28

  Query Match             6.8%;  Score 392;  DB 2;  Length 348;
  Best Local Similarity   26.5%;  Pred. No. 5.6e-27;
  Matches  116;  Conservative   56;  Mismatches  142;  Indels  124;  Gaps    6;

Qy        291 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 350

Db         28 NSQHSNKGESSASFLRAARAGNLDRVLE-LLRSGTDINTCNANGLNALHLASKEGHHEVV 86
```

```
Qy        351 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 410
              ::| :| |:|    | |:|| |:  |      :|:  | : |: || ||
Db         87 RELLKRKADVDAATRKGNTALHIASLAGQELIVTVLVENGANVNVQSLNGF--------- 137

Qy        411 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 470
                                                              |||: ||
Db        138 ------------------------------------------------TPLYMAAQE 146

Qy        471 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 530
              |  ||| ||| | |:     : |  ||  |   ||  |  :|:::     |    :   |
Db        147 NHESVVRYLLAHNANQALSTEDGFTPLAVALQQGHDRVVAVLLENDTRGKV----RLPAL 202

Qy        531 HEAAAKGKYEICKLLLQ--HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 588
              | || |   :   |||  | :| | |:
Db        203 HIAAKKDDTKAATLLLQNEHNSDVTSKS-------------------------------- 230

Qy        589 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 648
                                    ||||:|| | |  ||: ||: ||:|| | :  :
Db        231 ------------------------GFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNIS 266

Qy        649 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 708
              |||  |  :|  :: :||: : | ::   :   |||| |:: |  |:  ||| ||  : |
Db        267 PLHVATKWGRTNMVSLLLAHGAVIDCRTRDLLTPLHCASRSGHDQVVDLLLEKGAPISAK 326

Qy        709 NQEGQTPLDLVS-ADDVS 725
              : |  || : :  |||:
Db        327 TKNGLAPLHMAAQVDDVT 344
```

RESULT 43
US-09-065-474-28
; Sequence 28, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-065-474-28

Query Match 6.8%; Score 392; DB 3; Length 348;

Best Local Similarity 26.5%; Pred. No. 5.6e-27;
Matches 116; Conservative 56; Mismatches 142; Indels 124; Gaps 6;

```
Qy        291 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 350
              | :|    |::     |:      :: | ||| | :||       |  ||:||:: |::||
Db         28 NSQHSNKGESSASFLRAARAGNLDRVLE-LLRSGTDINTCNANGLNALHLASKEGHHEVV 86

Qy        351 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 410
              ::| :| |:|    | |:|| |:  |      :|:  | : |: || ||
Db         87 RELLKRKADVDAATRKGNTALHIASLAGQELIVTVLVENGANVNVQSLNGF--------- 137

Qy        411 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 470
                                                              |||: ||
Db        138 ------------------------------------------------TPLYMAAQE 146

Qy        471 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 530
              |  ||| ||| | |:     : |  ||  |   ||  |  :|:::     |    :   |
Db        147 NHESVVRYLLAHNANQALSTEDGFTPLAVALQQGHDRVVAVLLENDTRGKV----RLPAL 202

Qy        531 HEAAAKGKYEICKLLLQ--HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 588
              | || |   :   |||  | :| | |:
Db        203 HIAAKKDDTKAATLLLQNEHNSDVTSKS-------------------------------- 230

Qy        589 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 648
                                    ||||:|| | |  ||: ||: ||:|| | :  :
Db        231 ------------------------GFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNIS 266

Qy        649 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 708
              |||  |  :|  :: :||: : | ::   :   |||| |:: |  |:  ||| ||  : |
Db        267 PLHVATKWGRTNMVSLLLAHGAVIDCRTRDLLTPLHCASRSGHDQVVDLLLEKGAPISAK 326

Qy        709 NQEGQTPLDLVS-ADDVS 725
              : |  || : :  |||:
Db        327 TKNGLAPLHMAAQVDDVT 344
```

RESULT 44
US-09-557-034-28
; Sequence 28, Application US/09557034
; Patent No. 6365569
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; PROTEINS, NUCLEIC ACID MOLECULES, AND
; USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,034
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/065,474
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272

```
;             TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 28:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 348 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-557-034-28

  Query Match             6.8%;  Score 392;  DB 4;  Length 348;
  Best Local Similarity   26.5%;  Pred. No. 5.6e-27;
  Matches  116;  Conservative   56;  Mismatches  142;  Indels  124;  Gaps    6;

Qy        291 NFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHNDVV 350
              |  :|    |::     |:      :: | ||| | :||       |  ||:||:: |::||
Db         28 NSQHSNKGESSASFLRAARAGNLDRVLE-LLRSGTDINTCNANGLNALHLASKEGHHEVV 86

Qy        351 EVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNEN 410
               ::| :| |:|    | |:|| |:  |       :|:  | : |: || ||
Db         87 RELLKRKADVDAATRKGNTALHIASLAGQELIVTVLVENGANVNVQSLNGF--------- 137

Qy        411 VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGY 470
                                                        |||: ||
Db        138 ----------------------------------------------------TPLYMAAQE 146

Qy        471 NRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPL 530
              |  ||| ||| | |:     : |  ||  |   ||  |  :|:::     |    :  |
Db        147 NHESVVRYLLLAHNANQALSTEDGFTPLAVALQQGHDRVVAVLLENDTRGKV----RLPAL 202

Qy        531 HEAAAKGKYEICKLLLQ--HGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAK 588
              | || |   :   ||||  | :| | |:
Db        203 HIAAKKDDTKAATLLLQNEHNSDVTSKS------------------------------- 230

Qy        589 KGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLI 648
                                  ||||:|| | |  ||: ||: ||:|| | :  :
Db        231 -----------------------GFTPLHIAAHYGNENVAQLLLEKGANVNYQARHNIS 266

Qy        649 PLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK 708
              ||| |  :|  :: :||: : | ::   :   ||||:|:: |  |:  ||| ||  : |
Db        267 PLHVATKWGRTNMVSLLLAHGAVIDCRTRDLLTPLHCASRSGHDQVVDLLLEKGAPISAK 326

Qy        709 NQEGQTPLDLVS-ADDVS 725
               : |  || : :  |||:
Db        327 TKNGLAPLHMAAQVDDVT 344


RESULT 45
US-09-509-802-2
; Sequence 2, Application US/09509802
; Patent No. 6489130
; GENERAL INFORMATION:
;  APPLICANT:  Immunex Corp.
;  APPLICANT:  Bird, Timothy
;  APPLICANT:  Virca, G.D.
;  TITLE OF INVENTION:  DEATH ASSOCIATED KINASE CONTAINING ANKYRIN REPEATS (DAKAR)
;  FILE REFERENCE:  2889-US
;  CURRENT APPLICATION NUMBER: US/09/509,802
;  CURRENT FILING DATE:  2000-06-02
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 2
;  LENGTH: 786
;  TYPE: PRT
;  ORGANISM: Mus sp.
US-09-509-802-2

  Query Match             6.8%;  Score 390.5;  DB 4;  Length 786;
  Best Local Similarity   27.4%;  Pred. No. 3.3e-26;
  Matches  136;  Conservative   66;  Mismatches  176;  Indels  119;  Gaps   12;

Qy        228 SLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSL 287
              || ||:  :         |  || ||   |::         |: |    | :|: | |

Db        388 SLSLSFERE--------ASTGDLGPTDIQKKK---------LVDAIISGDTSRLMKILQP 430

Qy        288 EMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAHN 347
              : |:     :  : || |  :    :::  : ||   || |   ::  ||||:| |:
Db        431 QDVDLV-LDSSASLLHLAVEA---GQEECVKWLLLNNANPNLTNRKGSTPLHMAVERKGR 486

Qy        348 DVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFT----A 403
              :||::: :  ||| |   |:|| ||  |  : ||||      | :  :| |     |
Db        487 GIVELLLARKTSVNAKDEDQWTALHFAAQNGDEASTRLLLEKNASVNEVDFEGRTPMHVA 546

Qy        404 LQMGNEN-VQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQS- 461
              | | || |: ||: |: :|                                ::|: :
Db        547 CQHGQENIVRTLLRRGVDVG--------------------------------LQGKDAW 573

Qy        462 TPLHFAAGYNRVSVVEYLL-QHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN 520
              |||:||    : :|: |  | |  |:|:   |  ||| |   ||| || :|:   : ||
Db        574 LPLHYAAWQGHLPIVKLLAKQPGVSVNAQTLDGRTPLHLAAQRGHYRVARILIDLCSDVN 633

Qy        521 VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGD 580
              :  |   |||| ||  |     :|||  ||       :|
Db        634 ICSLQAQTPLHVAAETGHTSTARLLLHRGAGKEALTSEG---------------------- 672

Qy        581 AALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVN 640
                                              | |||||   :|   : |::  |||
Db        673 --------------------------------YTALHLAAQNGHLATVKLLIEEKADVM 699

Qy        641 AQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLA 700
              |:      || ||: || :|   |:  :  :: :|:    : || |||   :|     ||
Db        700 ARGPLNQTALHLAAARGHSEVVEELVSAD-LIDLSDEQGLSALHLAAQGRHSQTVETLLK 758

Qy        701 HGAD---PTLKNQEGQT 714
              |||     :|| | ||:
Db        759 HGAHINLQSLKFQGGQS 775


Search completed: March 25, 2004, 16:45:40
Job time : 30.7436 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 16:15:57 ; Search time 1358.1 Seconds
(without alignments)
11877.169 Million cell updates/sec

Title: US-09-843-159B-1
Perfect score: 3797
Sequence: 1 ctttgaagacactggatttc.....cctttgctgaaaaaaaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3797	100.0	3797	4	AAA91487	Aaa91487 Tankyrase
2	3797	100.0	3797	7	ABQ77066	Abq77066 Tankyrase
3	3393.4	89.4	3816	4	AAA91488	Aaa91488 Tankyrase
4	3393.4	89.4	3816	7	ABQ77067	Abq77067 Tankyrase
5	3388.8	89.2	4275	3	AAC66825	Aac66825 Human tan
6	3388.8	89.2	6019	4	AAC85294	Aac85294 Human SPA
7	3387.2	89.2	4127	5	AAF63926	Aaf63926 Human tan
8	3387.2	89.2	4406	5	AAF63930	Aaf63930 Human tan
9	3387.2	89.2	4992	5	AAF63953	Aaf63953 Human tan
10	3387.2	89.2	5002	5	AAF63952	Aaf63952 Human tan
11	3387.2	89.2	5075	6	ABZ11674	Abz11674 Human pol
12	3387.2	89.2	6018	7	AAD54031	Aad54031 Human col
13	3382.8	89.1	4512	4	AAD02578	Aad02578 Human tan
14	3297	86.8	3400	2	AAX25366	Aax25366 Human Grb
15	3296.2	86.8	3508	5	AAF63837	Aaf63837 Human tan
16	2969.6	78.2	4493	3	AAC66823	Aac66823 Human tan
17	2968.2	78.2	4296	3	AAC66824	Aac66824 Human tan
18	2774.6	73.1	5005	4	AAC89706	Aac89706 Human adu
19	2439.8	64.3	2971	5	AAF63919	Aaf63919 Human tan
20	2439.8	64.3	3353	5	AAF63917	Aaf63917 Human tan
21	2357.8	62.1	3799	5	AAF63918	Aaf63918 Human tan
22	1813.4	47.8	3984	5	AAF63838	Aaf63838 Human tan
23	1813.4	47.8	4134	3	AAZ29627	Aaz29627 Human tan
24	1813.4	47.8	4134	3	AAC66826	Aac66826 Human tan
25	1813.4	47.8	4134	6	ABK48629	Abk48629 Human tan
26	1794.8	47.3	3254	4	AAC85295	Aac85295 Mouse SPA
27	1533.8	40.4	4435	9	ADE53995	Ade53995 Human pro
28	1449.6	38.2	4491	3	AAZ29632	Aaz29632 Human tru
29	1449.6	38.2	4491	6	ABK48634	Abk48634 Truncated
30	1359.4	35.8	1568	5	AAF63925	Aaf63925 Human tan
31	1297.8	34.2	4657	3	AAZ29633	Aaz29633 Human tru
32	1297.8	34.2	4657	6	ABK48635	Abk48635 Truncated
33	1240.4	32.7	1466	4	ABA08564	Aba08564 Human Grb
34	1172.2	30.9	1242	4	AAC89701	Aac89701 Human adu
35	1168.4	30.8	3308	5	AAF63990	Aaf63990 Human tan
36	1119.4	29.5	1121	7	ACA57422	Aca57422 Human adi
37	1084.8	28.6	1103	5	AAF63973	Aaf63973 Human TRF
38	1076.8	28.4	3006	4	AAH16624	Aah16624 Human cDN
39	1008.8	26.6	5482	5	AAF63955	Aaf63955 Drosophil
40	1008.8	26.6	5484	4	ABL08155	Abl08155 Drosophil
41	994	26.2	2409	6	AAK99406	Aak99406 DNA of AP
42	990	26.1	1007	4	AAC89703	Aac89703 Human adu
43	794.4	20.9	797	9	ADE53994	Ade53994 Human pro
44	786	20.7	1344	3	AAC77578	Aac77578 Human ORF
45	784.2	20.7	838	4	AAH08231	Aah08231 Human cDN

ALIGNMENTS

RESULT 1
AAA91487
ID AAA91487 standard; DNA; 3797 BP.
XX
AC AAA91487;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 1 coding sequence.
XX
KW Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 404. .3706
FT /*tag= a
FT /partial
FT /product= "TaHo1"
FT /note= "Tankyrase homologue isotype 1; No start codon
FT given"
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR P-PSDB; AAY97748.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 13; Fig 1; 63pp; English.
XX
CC This sequence encodes the Tankyrase homologue isotype 1 (TaHo-1) protein

CC of the invention. The invention also relates to the TaHo-2 protein. The
CC TaHo proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as
CC vaccine, and for construction of hybridisation probes for mapping the
CC gene which encodes TaHo and for the genetic analysis of individuals with
CC genetic disorders. The TaHo proteins, and their coding sequences are
CC useful in screening assays
XX
SQ Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 U; 0 Other;

Query Match 100.0%; Score 3797; DB 4; Length 3797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTTGAAGACACTGGATTTCATACTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTTGAAGACACTGGATTTCATACTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60

Qy         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120

Qy        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180

Qy        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240

Qy        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300

Qy        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360

Qy        361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420

Qy        421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480

Qy        481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540

Qy        541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600

Qy        601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660

Qy        661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720

Qy        721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780

Qy        781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840

Qy        841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900

Qy        901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960

Qy        961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020

Qy       1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080

Qy       1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140

Qy       1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200

Qy       1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260

Qy       1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320

Qy       1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380

Qy       1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440

Qy       1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500

Qy       1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560

Qy       1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620

Qy       1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680

Qy       1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740

Qy       1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800

Qy       1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860

Qy       1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
```

```
Db       1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
Qy       1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
Db       1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
Qy       1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
Db       1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
Qy       2041 CCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Db       2041 CCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Qy       2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Db       2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Qy       2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Db       2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Qy       2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Db       2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Qy       2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Db       2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Qy       2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Db       2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Qy       2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Db       2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Qy       2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520
Db       2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520
Qy       2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580
Db       2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580
Qy       2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640
Db       2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640
Qy       2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Db       2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Qy       2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Db       2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Qy       2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820
Db       2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820
Qy       2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880
Db       2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880
Qy       2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940
Db       2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940
Qy       2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000
Db       2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000
Qy       3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Db       3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Qy       3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120
Db       3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120
Qy       3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180
Db       3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180
Qy       3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240
Db       3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240
Qy       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
Db       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
Qy       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
Db       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
Qy       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Db       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Qy       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
Db       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
Qy       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
Db       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
Qy       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
Db       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
Qy       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660
Db       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660
Qy       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720
Db       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720
Qy       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780
Db       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780
Qy       3781 TTGCTGAAAAAAAAAAA 3797
Db       3781 TTGCTGAAAAAAAAAAA 3797
```

RESULT 2
ABQ77066
ID ABQ77066 standard; DNA; 3797 BP.
XX
AC ABQ77066;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 1 TaHo-1 DNA from clone TH-1.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-1;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP; ds.
XX

OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 404. .3706
FT /*tag= a
FT /product= "TaHo-1"
FT /partial
FT /note= "No start codon given"
XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
DR P-PSDB; ABG73730.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Claim 13; Fig 1A-B; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence encodes the tankyrase H isoform 1, TaHo-1, isolated from clone
CC TH-1 which is described in the disclosure of the invention
XX
SQ Sequence 3797 BP; 1153 A; 732 C; 850 G; 1062 T; 0 U; 0 Other;

Query Match 100.0%; Score 3797; DB 7; Length 3797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTTGAAGACACTGGATTTCATACTTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTTGAAGACACTGGATTTCATACTTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60

Qy         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120

Qy        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180

Qy        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240

Qy        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300

Qy        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360

Qy        361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420

Qy        421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480

Qy        481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540

Qy        541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600

Qy        601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660

Qy        661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720

Qy        721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780

Qy        781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840

Qy        841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900

Qy        901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960

Qy        961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020

Qy       1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080

Qy       1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140

Qy       1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200

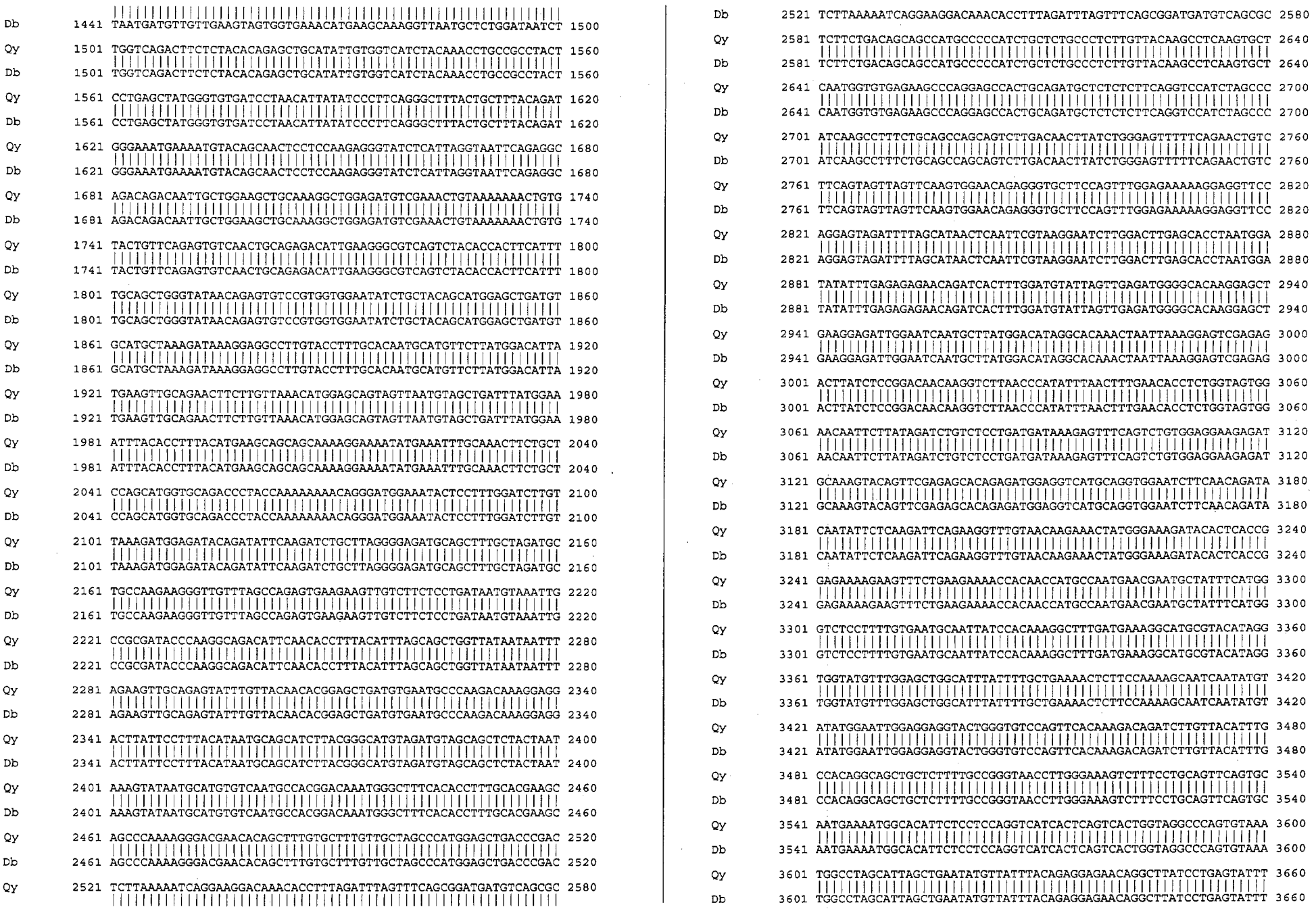
Qy       1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260

Qy       1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320

Qy       1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380

Qy       1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440

Qy       1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500
```


Db 1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500
Qy 1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560
Db 1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560
Qy 1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620
Db 1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620
Qy 1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
Db 1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
Qy 1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740
Db 1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740
Qy 1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800
Db 1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800
Qy 1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860
Db 1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860
Qy 1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
Db 1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
Qy 1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
Db 1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
Qy 1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
Db 1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
Qy 2041 CCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Db 2041 CCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Qy 2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Db 2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Qy 2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Db 2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Qy 2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Db 2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Qy 2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Db 2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Qy 2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Db 2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Qy 2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Db 2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Qy 2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520
Db 2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520
Qy 2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580
Db 2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580
Qy 2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640
Db 2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640
Qy 2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Db 2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Qy 2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Db 2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Qy 2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820
Db 2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820
Qy 2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880
Db 2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880
Qy 2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940
Db 2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940
Qy 2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000
Db 2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000
Qy 3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Db 3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Qy 3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120
Db 3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120
Qy 3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180
Db 3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180
Qy 3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240
Db 3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240
Qy 3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
Db 3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
Qy 3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
Db 3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
Qy 3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Db 3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Qy 3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
Db 3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
Qy 3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
Db 3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
Qy 3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
Db 3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
Qy 3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660
Db 3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660

```
Qy       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720

Qy       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780

Qy       3781 TTGCTGAAAAAAAAAAA 3797
             |||||||||||||||||
Db       3781 TTGCTGAAAAAAAAAAA 3797
```

RESULT 3
AAA91488
ID AAA91488 standard; DNA; 3816 BP.
XX
AC AAA91488;
XX
DT 06-AUG-2001 (first entry)
XX
DE Tankyrase homologue isotype 2 coding sequence.
XX
KW Tankyrase homologue isotype 2; TaHo-1; TaHo-2; cell proliferation;
KW cell cycle protein; cell cycle associated disorder; cancer; gene mapping;
KW chromosome mapping; gene therapy; vaccine; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2. .3726
FT /*tag= a
FT /partial
FT /product= "TaHo2"
FT /note= "Tankyrase homologue isotype 2; No start codon
FT given"
XX
PN WO200130987-A2.
XX
PD 03-MAY-2001.
XX
PF 25-OCT-2000; 2000WO-US041528.
XX
PR 25-OCT-1999; 99US-00427154.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B;
XX
DR WPI; 2001-300503/31.
DR P-PSDB; AAY97749.
XX
PT Novel recombinant cell cycle polypeptide, tankyrase H useful for inducing
PT or preventing cell proliferation in cells, and for diagnosing, treating
PT or preventing cell cycle associated disorders such as cancer.
XX
PS Claim 13; Fig 2; 63pp; English.
XX
CC This sequence encodes the Tankyrase homologue isotype 2 (TaHo-2) protein
CC of the invention. The invention also relates to the TaHo-1 protein. The
CC TaHo proteins are useful for inducing or preventing cell proliferation in
CC cells, and in the study or treatment of conditions mediated by the cell
CC cycle proteins, such as to diagnose, treat or prevent cell cycle
CC associated disorders, preferably cancer. The TaHo coding sequences are
CC useful as hybridisation probes, in chromosome and gene mapping and in the
CC generation of anti-sense DNA and RNA. The coding sequences are also
CC useful for the preparation of TaHo, for generating either transgenic
CC animals or knock out animals which, in turn, are useful in a development
CC and screening of therapeutically useful agents, in gene therapy, as
CC vaccine, and for construction of hybridisation probes for mapping the
CC gene which encodes TaHo and for the genetic analysis of individuals with
CC genetic disorders. The TaHo proteins, and their coding sequences are
CC useful in screening assays
XX
SQ Sequence 3816 BP; 1097 A; 806 C; 949 G; 964 T; 0 U; 0 Other;

Query Match 89.4%; Score 3393.4; DB 4; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141

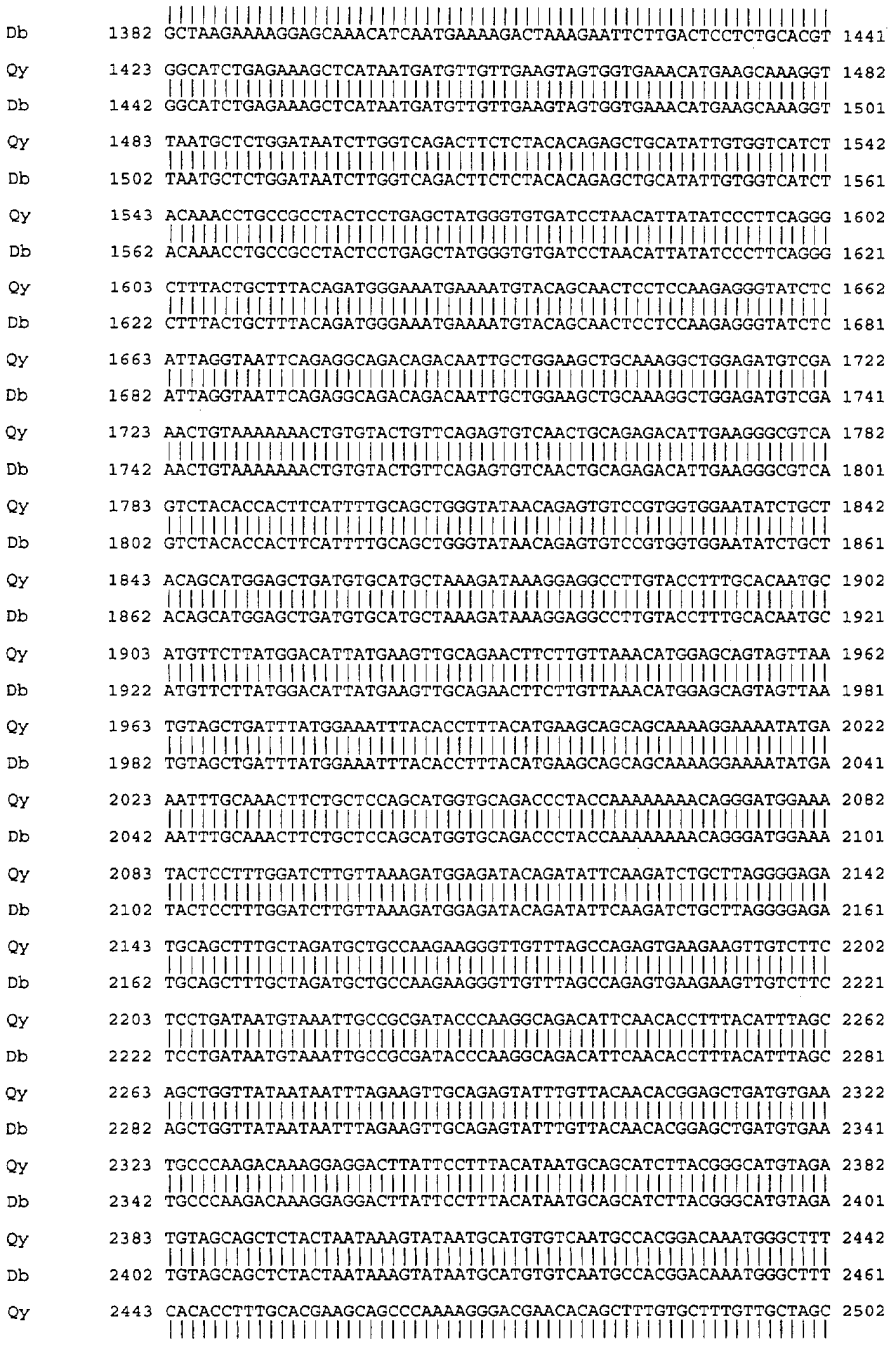
Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
```


Db 1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Qy 1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2101
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601

```
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
             ||||||||||||||||||||||||||||||||||||
Db       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAA 3816
```

RESULT 4
ABQ77067
ID ABQ77067 standard; DNA; 3816 BP.
XX
AC ABQ77067;
XX
DT 08-APR-2003 (first entry)
XX
DE Tankyrase H isotype 2 TaHo-2 DNA from clone K23.
XX
KW Tankyrase; tankyrase H; cell cycle protein; growth inhibitor; TaHo-2;
KW TaHo; cytostatic; immunostimulant; gene therapy; vaccine; cancer; tumour;
KW cell cycle related disorder; poly ADP-ribose polymerase; PARP; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 3. .3725
FT /*tag= a
FT /product= "TaHo-2"
FT /partial
FT /note= "no start codon"
XX
PN WO200286170-A1.
XX
PD 31-OCT-2002.
XX
PF 25-APR-2002; 2002WO-US013185.
XX
PR 25-APR-2001; 2001US-00843159.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Luo Y, Chan E, Xu X, Huang B, Ossovskaya V;
XX
DR WPI; 2003-093158/08.
DR P-PSDB; ABG73731.
XX
PT New recombinant nucleic acid encoding a cell cycle protein, useful for
PT diagnosing and treating a cell cycle related disorder, e.g. cancer.
XX
PS Claim 13; Fig 2A-B; 90pp; English.
XX
CC This invention describes a novel recombinant nucleic acid encoding the
CC cell cycle protein TaHo, a tankyrase H isoform. The products of the
CC invention have cytostatic and immunostimulant activity and can be used
CC for gene therapy and in vaccines. The cell cycle protein TaHo and the
CC nucleic acid encoding the protein are useful for diagnosing and treating
CC a cell cycle related disorder, e.g. cancer, or for inhibiting growth of a
CC tumour cell. The methods are useful for screening for a bioactive agent
CC capable of binding to a cell cycle protein tankyrase H, or a bioactive
CC agent capable of modulating a cell cycle protein tankyrase H or PARP
CC (poly ADP-ribose polymerase) activity. The antibodies are useful for
CC inducing an immune response against the cell cycle proteins. This
CC sequence encodes the tankyrase H isoform 2, TaHo-2, isolated from clone
CC K23 which is described in the disclosure of the invention
XX
SQ Sequence 3816 BP; 1097 A; 806 C; 949 G; 964 T; 0 U; 0 Other;

Query Match 89.4%; Score 3393.4; DB 7; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141

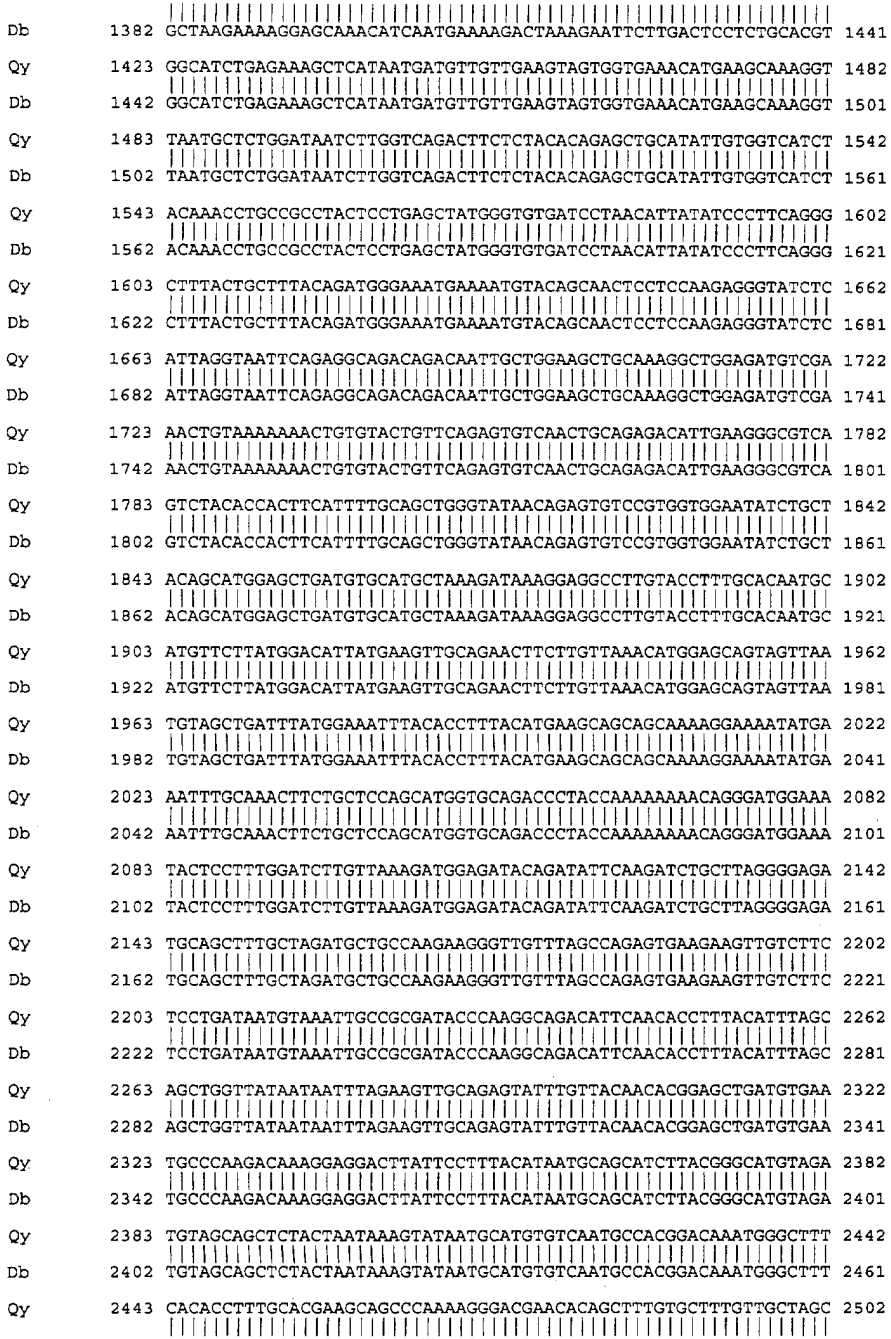
Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
```


Db 1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2101
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601

```
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
             ||||||||||||||||||||||||||||||||||||
Db       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
```

RESULT 5
AAC66825
ID AAC66825 standard; cDNA; 4275 BP.
XX
AC AAC66825;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase II coding sequence SEQ ID NO: 5.
XX
KW Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 284. .3784
FT /*tag= a
FT /product= "tankyrase II"
FT /transl_except= (pos:655. .660,aa:TyrThr)
XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009558.
XX
PR 09-APR-1999; 99US-0128577P.
PR 13-APR-1999; 99US-0129123P.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
DR P-PSDB; AAB27211.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 1; Fig 4; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase II coding
CC sequence. Its protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell
XX
SQ Sequence 4275 BP; 1247 A; 889 C; 1025 G; 1114 T; 0 U; 0 Other;

Query Match 89.2%; Score 3388.8; DB 3; Length 4275;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        481 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 540

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 600

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 660

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 720

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 780

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 840

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 900

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 960

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1020

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1080

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1140

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1200

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1260

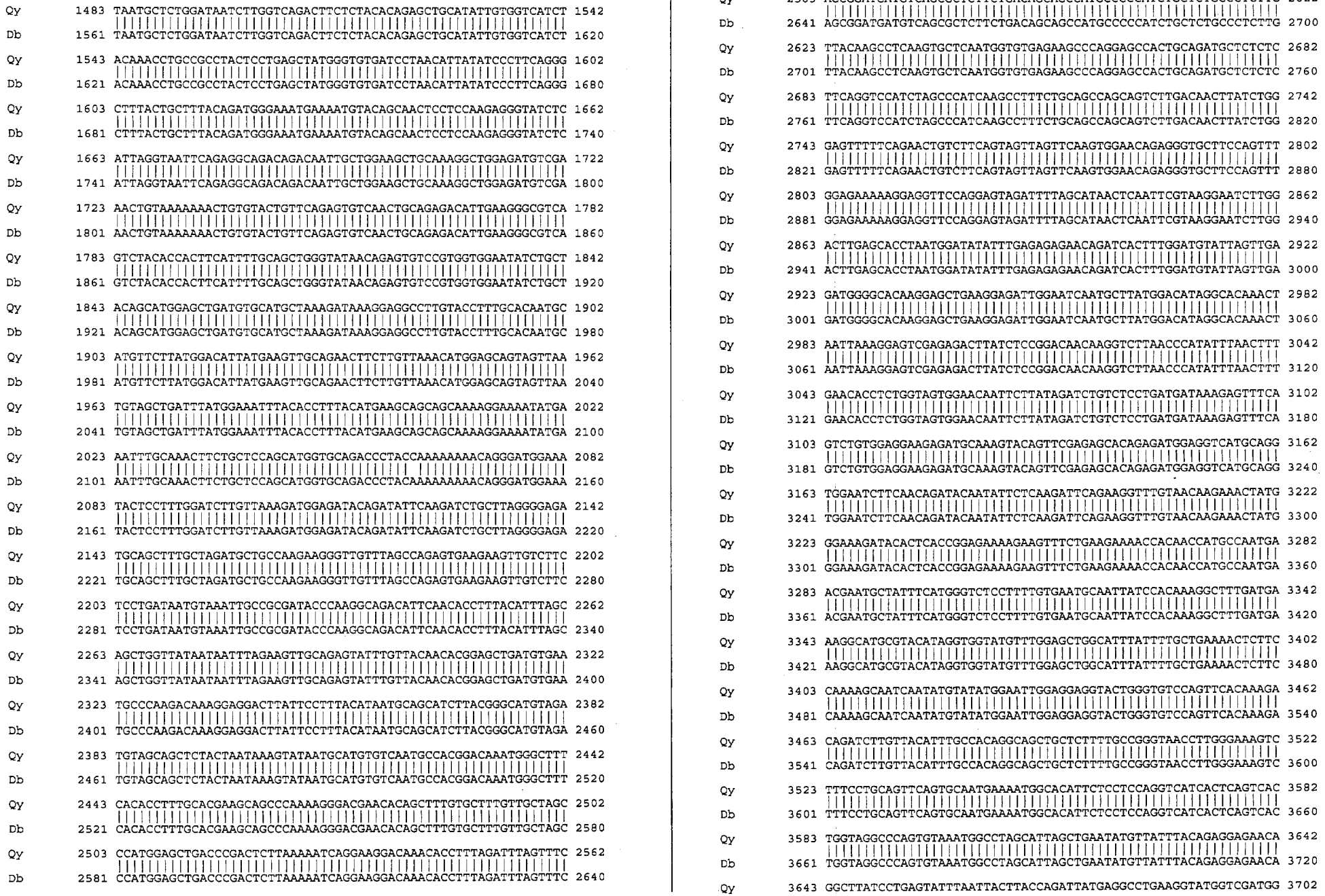
Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1320

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1380

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1440

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1500

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1560
```



```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3780

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3840

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
                ||||||||||||||||||||||||||||||||
Db       3841 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 6
AAC85294
ID AAC85294 standard; cDNA; 6019 BP.
XX
AC AAC85294;
XX
DT 29-MAR-2001 (first entry)
XX
DE Human SPANK cDNA.
XX
KW SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW adipocyte; insulin signalling pathway; hyperlipidaemia;
KW glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW cardiac insufficiency; coronary insufficiency; stroke;
KW high blood pressure; non-insulin dependent diabetes; hypertension;
KW hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   252. .3735
FT   /*tag= a
FT   /product= "Human SPANK"
FT   /transl_except= (pos:768. .878,aa:173. .209)
FT   /transl_except= (pos:1230. .1337,aa:327. .362)
FT   /transl_except= (pos:1374. .1445,aa:375. .398)
FT   /transl_except= (pos:1695. .1823,aa:482. .524)
FT   /transl_except= (pos:2172. .2282,aa:641. .677)
XX
PN WO200077225-A1.
XX
PD 21-DEC-2000.
XX
PF 09-JUN-2000; 2000WO-US015926.
XX
PR 11-JUN-1999; 99US-0138957P.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Chi N, Lodish HF;
XX
DR WPI; 2001-091404/10.
DR P-PSDB; AAB47022.
XX
PT New insulin signaling protein SPANK, useful for reducing body mass,
PT glucose intolerance or insulin resistance and for preventing or treating
PT obesity-related and muscle-related diseases.
XX
PS Claim 3; Fig 4; 65pp; English.
XX
CC This sequence represents the human SPANK cDNA. The SPANK protein
CC comprises 3 domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP
CC (poly adenosine diphosphate-ribose polymerase) catalytic domain; and (c)
CC an ANK domain composed of ankyrin repeats. SPANK is a cytosolic protein
CC which can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC perinuclear region of adipocytes. It is an effector in the insulin
CC signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC mass, reducing glucose intolerance or insulin resistance, for preventing
CC or treating obesity-related diseases or disorders, such as obesity,
CC cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC also useful for preventing or treating muscle-related diseases or
CC disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC Antibodies immunospecific for SPANK are useful for detecting the presence
CC of SPANK polypeptide in a biological sample
XX
SQ Sequence 6019 BP; 1732 A; 1173 C; 1341 G; 1773 T; 0 U; 0 Other;

Query Match 89.2%; Score 3388.8; DB 4; Length 6019;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 508

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 568

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 628

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 688

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 748

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 808

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 868

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 928

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 988

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1048

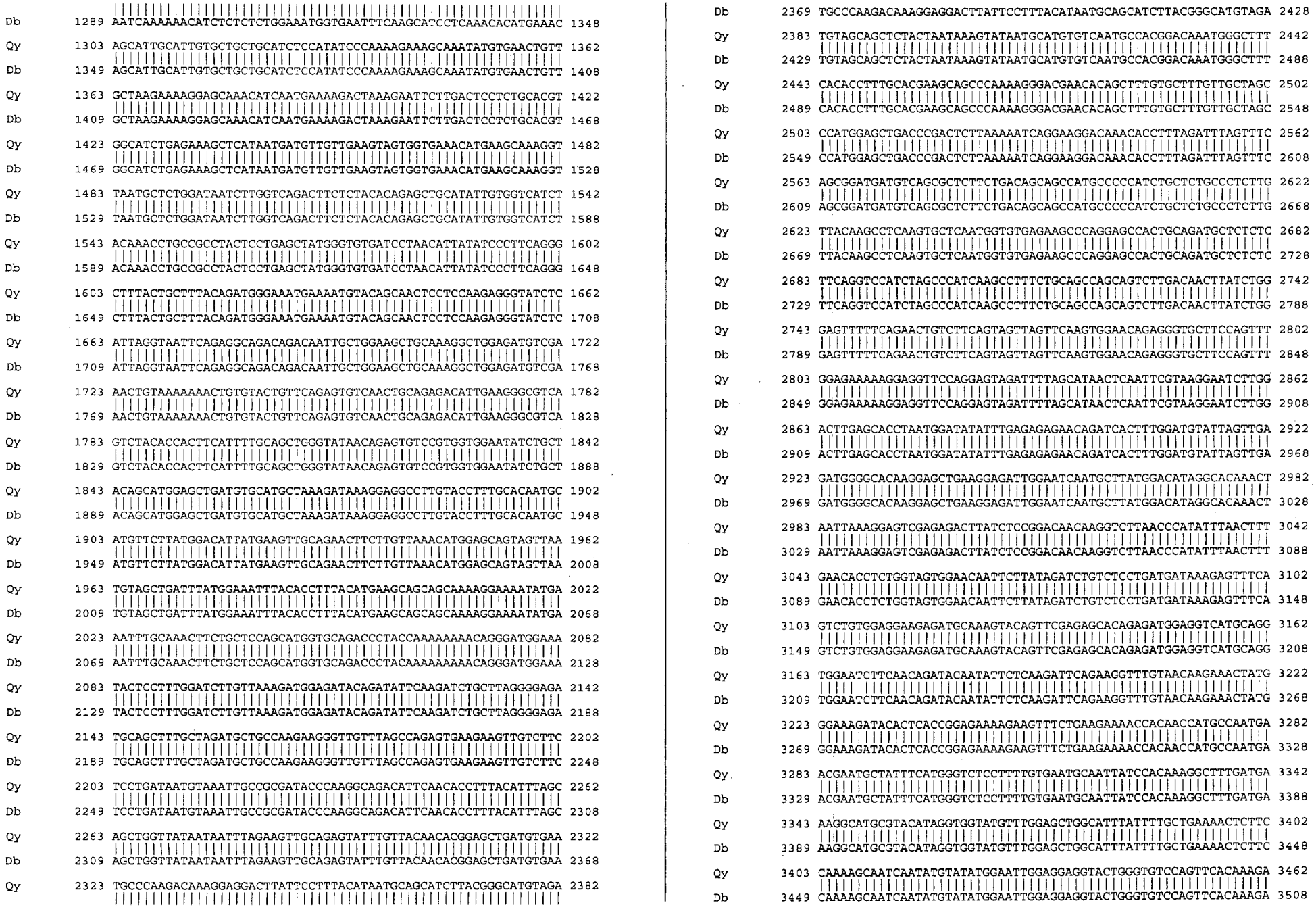
Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1108

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1168

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1228

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1288

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
```


Db 1289 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1348
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1349 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1408
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1409 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1468
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1469 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1528
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1529 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1588
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1589 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1648
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1649 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1708
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1709 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1768
Qy 1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1769 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1828
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1829 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1888
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1889 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1948
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1949 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2008
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2009 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2068
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 2069 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2128
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2129 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2188
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2189 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2248
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2249 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2308
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2309 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2368
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2369 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2428
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2429 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2488
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2489 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2548
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2549 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2608
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2609 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2668
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2669 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2728
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2729 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2788
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2789 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2848
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2849 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2908
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2909 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2968
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2969 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3028
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3029 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3088
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3089 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3148
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3149 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3208
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3209 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3268
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3269 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3328
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3329 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3388
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3389 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3448
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3449 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3508

```
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3568

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3628

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3688

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3748

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3808

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3809 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3840
```

RESULT 7
AAF63926
ID AAF63926 standard; DNA; 4127 BP.
XX
AC AAF63926;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 100.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66288.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 158-162; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 4127 BP; 1245 A; 813 C; 934 G; 1135 T; 0 U; 0 Other;

Query Match 89.2%; Score 3387.2; DB 5; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

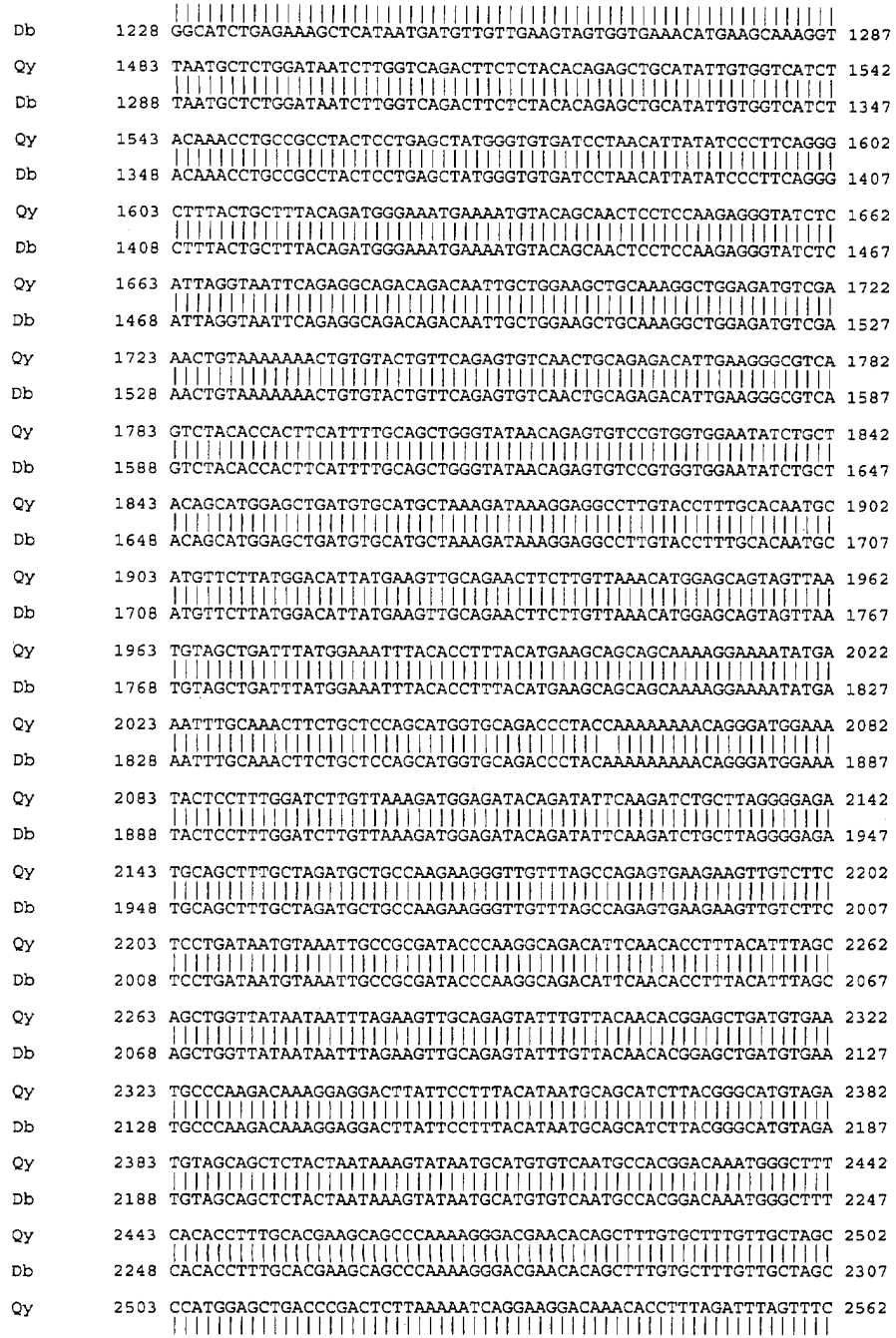
Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

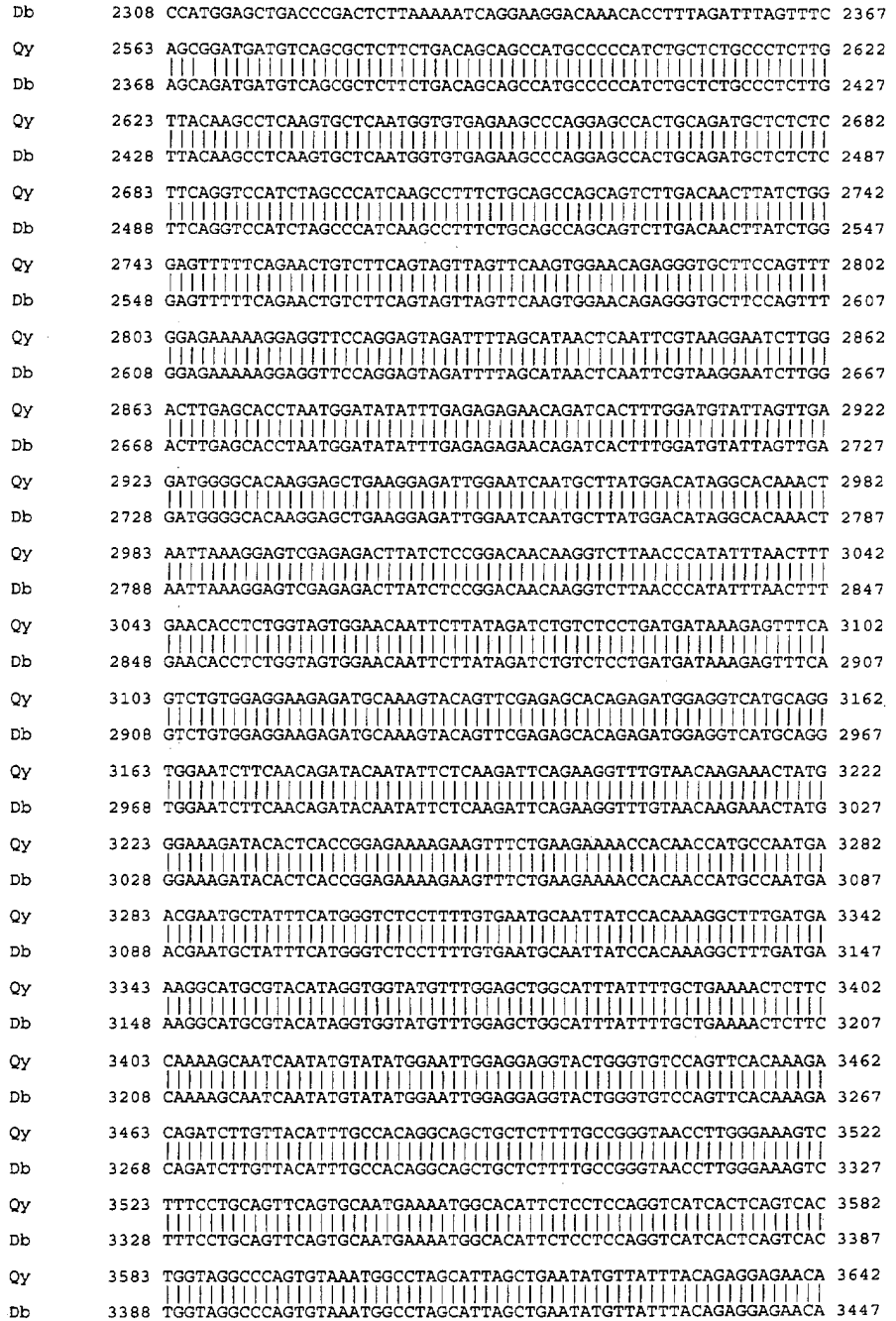
Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
```


Db 1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527
Qy 1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1587
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1588 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1647
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1648 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1707
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1708 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1767
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1768 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1827
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1887
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1888 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1947
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1948 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2007
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2008 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2067
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2068 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2127
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2188 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2247
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2248 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2307
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562


Db 2308 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2427
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2428 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2487
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2488 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2547
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2607
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2608 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2728 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2787
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2847
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 2848 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2907
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 2908 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2967
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 2968 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3027
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3028 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3087
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3088 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3207
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3208 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3267
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3268 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3327
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3328 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3387
Qy 3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db 3388 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3447

```
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3507

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3567

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3568 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3599


RESULT 8
AAF63930
ID   AAF63930 standard; DNA; 4406 BP.
XX
AC   AAF63930;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 clone consensus oligonucleotide SEQ ID NO: 106.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66290.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 2; Page 168-173; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 4406 BP; 1268 A; 910 C; 1054 G; 1174 T; 0 U; 0 Other;

  Query Match             89.2%;  Score 3387.2;  DB 5;  Length 4406;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        487 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 546

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 606

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 666

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 726

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 786

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 846

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 906

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 966

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1026

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1086

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1146

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1206

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1266

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1326

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1386

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1446

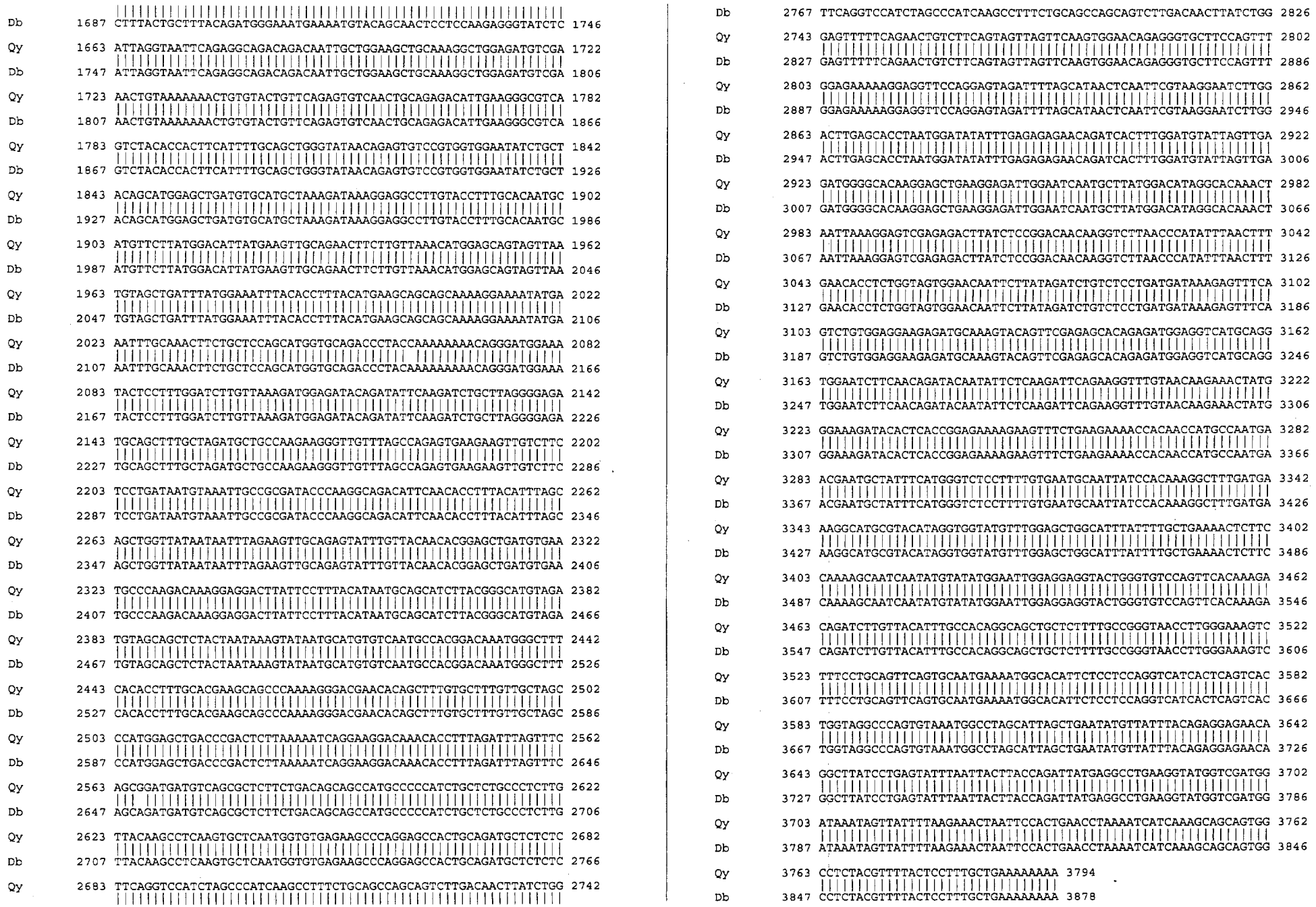
Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1506

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1566

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1626

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1686

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
```


Db 1687 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1746
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1747 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1806
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1807 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1866
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1867 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1926
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1927 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1986
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1987 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2046
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2047 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2106
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2107 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2166
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2167 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2226
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2227 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2286
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2287 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2346
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2347 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2406
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2407 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2466
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2467 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2526
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2527 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2586
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2587 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2646
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2647 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2706
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2707 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2766
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2767 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2826
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2827 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2886
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2887 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2946
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2947 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3006
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 3007 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3066
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3067 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3126
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3127 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3186
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3187 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3246
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3247 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3306
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3307 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3366
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3367 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3426
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3427 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3486
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3487 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3546
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3547 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3606
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3607 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3666
Qy 3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db 3667 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3726
Qy 3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db 3727 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3786
Qy 3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db 3787 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3846
Qy 3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db 3847 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3878

RESULT 9
AAF63953
ID AAF63953 standard; DNA; 4992 BP.
XX
AC AAF63953;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-SHORT coding sequence SEQ ID NO: 134.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66295.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Claim 6; Page 195-200; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 4992 BP; 1454 A; 1064 C; 1191 G; 1283 T; 0 U; 0 Other;

Query Match 89.2%; Score 3387.2; DB 5; Length 4992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 1132

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 1192

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 1252

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 1312

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 1372

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1432

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1492

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1552

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1612

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1672

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1732

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1792

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1852

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1912

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1972

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 2032

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 2092

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 2152

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 2212

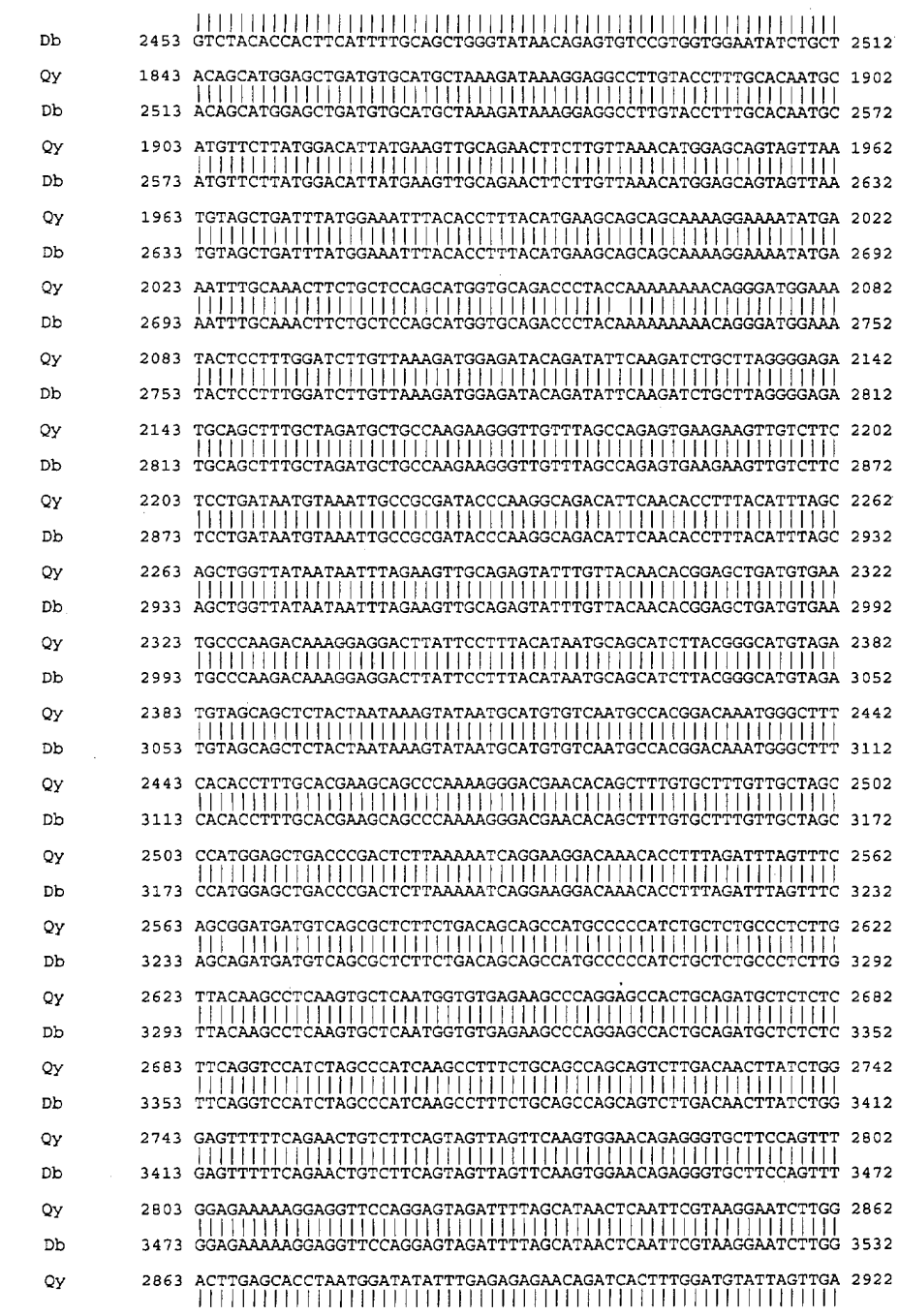
Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 2272

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 2332

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 2392

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2452

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
```



RESULT 10
AAF63952
ID AAF63952 standard; DNA; 5002 BP.
XX
AC AAF63952;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 TANK2-LONG coding sequence SEQ ID NO: 132.
XX

KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66294.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Claim 5; Page 185-190; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 5002 BP; 1456 A; 1065 C; 1196 G; 1285 T; 0 U; 0 Other;

Query Match 89.2%; Score 3387.2; DB 5; Length 5002;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      1083 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 1142

Qy       463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1143 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 1202

Qy       523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1203 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 1262

Qy       583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1263 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 1322

Qy       643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1323 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 1382

Qy       703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1383 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1442

Qy       763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1443 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1502

Qy       823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1503 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1562

Qy       883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1563 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1622

Qy       943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1623 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1682

Qy      1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1683 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1742

Qy      1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1743 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1802

Qy      1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1803 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1862

Qy      1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1863 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1922

Qy      1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1923 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1982

Qy      1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1983 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 2042

Qy      1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2043 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 2102

Qy      1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2103 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 2162

Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2163 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 2222

Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2223 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 2282

Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2283 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 2342

Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2343 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 2402

Qy      1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2403 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2462

Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2463 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2522

Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2523 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 2582

Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2583 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2642

Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
```

```
Db       2643 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2702
Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db       2703 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2762
Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       2763 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2822
Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       2823 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2882
Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2883 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2942
Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2943 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 3002
Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       3003 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 3062
Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       3063 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 3122
Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       3123 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 3182
Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       3183 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 3242
Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       3243 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 3302
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       3303 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 3362
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       3363 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 3422
Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       3423 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 3482
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       3483 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3542
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       3543 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3602
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       3603 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3662
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       3663 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3722
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       3723 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3782
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       3783 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3842
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       3843 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3902
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3903 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3962
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3963 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 4022
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       4023 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 4082
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       4083 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 4142
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       4143 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 4202
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       4203 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 4262
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       4263 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 4322
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       4323 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 4382
Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       4383 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 4442
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       4443 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4474
```

RESULT 11
ABZ11674
ID ABZ11674 standard; cDNA; 5075 BP.
XX
AC ABZ11674;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 556.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200270539-A2.

XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR P-PSDB; ABP69457.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 1; SEQ ID NO 556; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5075 BP; 1487 A; 1055 C; 1166 G; 1367 T; 0 U; 0 Other;

Query Match 89.2%; Score 3387.2; DB 6; Length 5075;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db   450 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAAAATGGTGCAAATGTCCAAGC 509

Qy   463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db   510 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 569

Qy   523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db   570 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 629

Qy   583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db   630 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 689

Qy   643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db   690 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 749

Qy   703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db   750 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 809

Qy   763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db   810 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 869

Qy   823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db   870 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 929

Qy   883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db   930 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 989

Qy   943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db   990 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1049

Qy  1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db  1050 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1109

Qy  1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db  1110 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1169

Qy  1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db  1170 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1229

Qy  1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db  1230 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1289

Qy  1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db  1290 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1349

Qy  1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db  1350 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1409

Qy  1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db  1410 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1469

Qy  1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db  1470 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1529

Qy  1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db  1530 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1589

Qy  1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db  1590 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1649

Qy  1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db  1650 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1709

Qy  1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db  1710 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1769

Qy  1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db  1770 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1829

Qy  1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db  1830 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1889

Qy  1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db  1890 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1949
```

```
Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1950 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2009

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2010 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2069

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
             |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       2070 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2129

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2189

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2249

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2250 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2309

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2310 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2369

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2370 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2429

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2430 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2489

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2490 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2549

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2550 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2609

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2610 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2669

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2670 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2729

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2730 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2789

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2849

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2909

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2969

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3029

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3089

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3149

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3209

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3269

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3329

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3389

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3449

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3450 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3509

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3569

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3629

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3689

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3749

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3750 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3809

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3810 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3841
```

RESULT 12
AAD54031
ID AAD54031 standard; DNA; 6018 BP.
XX
AC AAD54031;
XX
DT 17-JUN-2003 (first entry)
XX
DE Human colon cancer-associated polypeptide gene, TNKL (BC-203).
XX
KW Human; colon cancer-associated polypeptide; immune response; therapy;
KW colon cancer; gene; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   33. .3283

```
FT                   /*tag=  a
FT                   /product= "Human colon cancer-associated polypeptide"
FT                   /transl_except= (pos:84. .87, aa:Ser)
FT                   /transl_except= (pos:262. .265, aa:Asp)
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200290986-A1.
XX
PD   14-NOV-2002.
XX
PF   02-MAY-2002; 2002WO-US013994.
XX
PR   04-MAY-2001; 2001US-00849602.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chen Y,  Old LJ,  Scanlan MJ,  Stockert E;
XX
DR   WPI; 2003-112003/10.
DR   P-PSDB; AAE35352.
XX
PT   Diagnosing colon cancer in a subject comprises identifying colon cancer-
PT   associated polypeptides as antigens that elicit immune response in colon
PT   cancer.
XX
PS   Claim 1; Page 85-89; 122pp; English.
XX
CC   The invention relates to a method for diagnosing colon cancer in a
CC   subject which comprises identifying colon cancer-associated polypeptides
CC   as antigens that elicit immune response in colon cancer. The method is
CC   useful for diagnosing, determining onset, progression, or regression of
CC   colon cancer in a subject, or for selecting a course of treatment of a
CC   subject having or suspected of having colon cancer. The colon cancer-
CC   associated polypeptides are useful as markers for diagnosing colon
CC   cancer, and for following the course of treatment of colon cancer. The
CC   present sequence is human colon cancer-associated polypeptide gene
XX
SQ   Sequence 6018 BP; 1741 A; 1171 C; 1328 G; 1778 T; 0 U; 0 Other;

  Query Match             89.2%;  Score 3387.2;  DB 7;  Length 6018;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        418 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 477

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 537

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 597

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 657

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 717

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 777

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 837

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 897

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 957

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1017

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1077

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1137

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1197

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1257

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1317

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1377

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1437

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1497

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1557

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1617

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1677

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1737

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1797

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1857

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1858 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1917

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
```

```
Db       1918 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1977
Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       1978 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2037
Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db       2038 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2097
Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       2098 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2157
Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       2158 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2217
Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2218 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2277
Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2278 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2337
Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2338 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2397
Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2398 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2457
Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2458 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2517
Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2518 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2577
Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2578 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2637
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2638 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2697
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2698 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2757
Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2758 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2817
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       2818 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2877
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2878 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2937
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2938 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2997
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2998 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3057
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       3058 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3117
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       3118 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3177
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       3178 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3237
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3238 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3297
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3298 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3357
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3358 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3417
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3418 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3477
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3478 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3537
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3538 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3597
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3598 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3657
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3658 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3717
Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       3718 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3777
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       3778 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3809
```

```
RESULT 13
AAD02578
ID   AAD02578 standard; DNA; 4512 BP.
XX
AC   AAD02578;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human tankyrase homolog protein (THP) DNA.
XX
KW   Human; tankyrase homolog protein; THP; gene therapy; cancer; tumour;
KW   basal cell carcinoma; therapy; genetic mapping; cytostatic; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             23. .3523
FT                   /*tag= a
```

```
FT                  /product= "Human tankyrase homolog protein (THP)"
FT                  /note= "This region is specifically claimed in claim 4 as
FT                  SEQ ID NO:4"
XX
PN   WO200104326-A1.
XX
PD   18-JAN-2001.
XX
PF   03-JUL-2000; 2000WO-EP006609.
XX
PR   09-JUL-1999;   99US-00350982.
XX
PA   (PHAA ) PHARMACIA & UPJOHN SPA.
XX
PI   Berthelsen J,  Toma S,  Isacchi A;
XX
DR   WPI; 2001-168422/17.
DR   P-PSDB; AAY72589.
XX
PT   New tankyrase homolog protein (THP) polynucleotide and polypeptide useful
PT   in gene therapy, diagnosis and treatment or prevention of unregulated
PT   cell growth, such as cancer or tumor cell growth.
XX
PS   Claim 1; Page 41-42; 60pp; English.
XX
CC   The present sequence is a DNA coding for human tankyrase homolog protein
CC   (THP). The THP polypeptides and polynucleotides of the invention are
CC   useful in gene therapy and for treating or preventing unregulated cell
CC   growth such as cancer or tumour (e.g. basal cell carcinoma). The nucleic
CC   acid molecules of the invention and their fragments are useful for
CC   restriction fragment length polymorphism (RFLP) associated with certain
CC   disorders, as well as for genetic mapping. Antisense oligonucleotides, or
CC   fragments of nucleic acid encoding THP are useful as diagnostic tools for
CC   probing the expression of Thp gene in various tissues. THP can be used as
CC   antigens for raising antibodies against them and in assays for
CC   identifying compounds that modulate their activity. They are used in the
CC   manufacture of a medicament directed towards cancers or tumours. THP are
CC   also useful for screening compounds in a variety of drugs screening
CC   techniques and as a research tool for identification, characterisation
CC   and purification of interacting, regulatory proteins
XX
SQ   Sequence 4512 BP; 1356 A; 878 C; 1002 G; 1268 T; 0 U; 8 Other;
```

Query Match 89.1%; Score 3382.8; DB 4; Length 4512;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3381; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db        220 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 279

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db        280 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 339

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db        340 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 399

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db        400 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 459

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db        460 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 519

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db        520 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 579

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db        580 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 639

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db        640 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 699

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db        700 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 759

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db        760 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 819

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db        820 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 879

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db        880 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 939

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db        940 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 999

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db       1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1059

Qy       1243 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db       1060 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1119

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db       1120 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1179

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db       1180 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1239

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db       1240 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1299

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db       1300 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1359

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db       1360 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1419

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db       1420 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1479

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db       1480 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1539

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db       1540 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1599

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db       1600 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1659

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db       1660 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1719
```

```
Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       1720 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1779

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       1780 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1839

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db       1840 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1899

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       1900 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1959

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       1960 TGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2019

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2020 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2079

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2080 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2139

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2140 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2199

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2200 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2259

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2260 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGKTGCTAGC 2319

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2320 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2379

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2380 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2439

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2440 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2499

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2500 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2559

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2560 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2619

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       2620 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGNATCTTGG 2679

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2680 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2739

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2740 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2799

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2800 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2859

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       2860 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2919

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       2920 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2979

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       2980 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3039

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3040 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3099

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3100 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3159

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3160 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3219

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3220 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3279

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3280 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3339

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3340 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3399

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3400 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3459

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3460 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3519

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       3520 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3579

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       3580 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3611
```

RESULT 14
AAX25366
ID AAX25366 standard; cDNA; 3400 BP.
XX
AC AAX25366;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human Grb7 effector 2.2412 cDNA.
XX
KW Grb7 effector; 2.2412 protein; human; signal transduction; tumour marker;
KW breast cancer; prostate cancer; prognosis; diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers

```
FT   CDS             1. .3225
FT                   /*tag=  a
XX
PN   WO9915647-A1.
XX
PD   01-APR-1999.
XX
PF   23-SEP-1998;   98WO-AU000795.
XX
PR   23-SEP-1997;   97AU-00009388.
XX
PA   (GARV-) GARVAN INST MEDICAL RES.
XX
PI   Daly RJ,  Sutherland RL;
XX
DR   WPI; 1999-254707/21.
DR   P-PSDB; AAY05734.
XX
PT   New candidate effector for the Grb7 family of signaling proteins, and
PT   specific antibody, useful for detection and treatment of cancer.
XX
PS   Claim 4; Fig 1; 24pp; English.
XX
CC   This is the nucleotide sequence of an isolated polynucleotide molecule
CC   encoding a candidate effector protein, termed 2.2412 (see AAY05734), for
CC   the Grb7 family of signalling proteins. A partial clone was isolated from
CC   a human liver cDNA library using a yeast two-hybrid system assay with
CC   Grb14 as bait. This clone was used to screen a human placental cDNA
CC   library that provided the C-terminal end of the 2.2412 sequence. A 5'
CC   sequence for the clone, including the initiation codon, was not
CC   identified. Analysis of the sequence revealed significant homology to a
CC   large number of proteins containing ankyrin-like repeats. The gene was
CC   localised to between chromosome 10q23.2 and proximal 10q23.32. Deletions
CC   in the 10q22-25 region have been detected in human breast, prostate,
CC   renal, small cell lung and endometrial carcinomas, glioblastoma
CC   multiforme, melanoma and meningiomas. Detection of the protein encoded by
CC   the 2.2412 cDNA in a sample should provide a useful tumour marker and/or
CC   prognostic indicator for certain human cancers, in particular breast
CC   cancer and prostate cancer. Antagonism of the interaction between Grb7
CC   family members and the encoded protein should provide a novel treatment
CC   strategy for human diseases exhibiting aberrant receptor tyrosine kinase
CC   signalling, such as cancer. Oligonucleotide probes can be used in methods
CC   of detecting the presence of 2.2412 mRNA in a sample
XX
SQ   Sequence 3400 BP; 1059 A; 665 C; 760 G; 916 T; 0 U; 0 Other;

  Query Match           86.8%;  Score 3297;  DB 2;  Length 3400;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       482 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 541
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60

Qy       542 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120

Qy       602 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180

Qy       662 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240

Qy       722 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300

Qy       782 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 841
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360

Qy       842 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420

Qy       902 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 961
            ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480

Qy       962 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1021
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540

Qy      1022 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1081
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600

Qy      1082 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1141
            |||||||||||||||||||||||||||||||||||||||||||||||| | |||||||||
Db       601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660

Qy      1142 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1201
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720

Qy      1202 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1261
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780

Qy      1262 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840

Qy      1322 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1381
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900

Qy      1382 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1441
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960

Qy      1442 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1501
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020

Qy      1502 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1561
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080

Qy      1562 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy      1622 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1681
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200

Qy      1682 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1741
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260

Qy      1742 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1801
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320

Qy      1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1861
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380

Qy      1862 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1921
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440
```

```
Qy       1922 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1981
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500

Qy       1982 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2041
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560

Qy       2042 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2101
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620

Qy       2102 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680

Qy       2162 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740

Qy       2222 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2281
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800

Qy       2282 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860

Qy       2342 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920

Qy       2402 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2461
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980

Qy       2462 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040

Qy       2522 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2581
             |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db       2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100

Qy       2582 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2641
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160

Qy       2642 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2701
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220

Qy       2702 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2761
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280

Qy       2762 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2821
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340

Qy       2822 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2881
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400

Qy       2882 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2941
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460

Qy       2942 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3001
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520

Qy       3002 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3061
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580

Qy       3062 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640

Qy       3122 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700

Qy       3182 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3242 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820

Qy       3302 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880

Qy       3362 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940

Qy       3422 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000

Qy       3482 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060

Qy       3542 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120

Qy       3602 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180

Qy       3662 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240

Qy       3722 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300

Qy       3782 TGCTGAAAAAAAA 3794
             |||||||||||||
Db       3301 TGCTGAAAAAAAA 3313
```

RESULT 15
AAF63837
ID AAF63837 standard; DNA; 3508 BP.
XX
AC AAF63837;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 related coding sequence SEQ ID NO: 1.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX

PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66278.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Disclosure; Page 105-109; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 3508 BP; 1046 A; 710 C; 843 G; 909 T; 0 U; 0 Other;

Query Match 86.8%; Score 3296.2; DB 5; Length 3508;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3298; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527

Qy       1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1587

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1647

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1707

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1708 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1767

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1768 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1827

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1887
```

```
Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       1888 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1947

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       1948 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2007

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2008 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2067

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2068 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2127

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2188 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2247

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2248 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2307

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2308 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2427

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2428 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2487

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2488 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2547

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2607

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       2608 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2728 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2787

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2847

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       2848 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2907

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       2908 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2967

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       2968 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3027

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3028 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3087

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3088 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3207

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3208 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3267

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3268 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3327

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3328 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3387

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3388 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3447

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3448 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3507

Qy       3703 A 3703
              |
Db       3508 A 3508
```

RESULT 16
AAC66823
ID AAC66823 standard; cDNA; 4493 BP.
XX
AC AAC66823;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase II coding sequence SEQ ID NO: 1.
XX
KW Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	1. .4002
FT		/*tag= a
FT		/transl_except= (pos:595. .597,aa:Xaa)
FT		/transl_except= (pos:625. .630,aa:XaaXaa)
FT		/transl_except= (pos:637. .639,aa:Xaa)
FT		/transl_except= (pos:694. .696,aa:Xaa)
FT		/transl_except= (pos:724. .726,aa:Xaa)
FT		/transl_except= (pos:730. .732,aa:Xaa)
FT		/transl_except= (pos:763. .765,aa:Xaa)
FT		/transl_except= (pos:790. .792,aa:Xaa)
FT		/transl_except= (pos:2176. .2178,aa:Xaa)
FT		/transl_except= (pos:2380. .2382,aa:Xaa)
FT		/transl_except= (pos:2398. .2400,aa:Xaa)
FT		/transl_except= (pos:2464. .2469,aa:XaaXaa)
FT		/transl_except= (pos:3163. .3165,aa:Xaa)
FT		/note= "Xaa=unknown"
FT		/product= "tankyrase II"
FT		/partial

XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009558.
XX
PR 09-APR-1999; 99US-0128577P.
PR 13-APR-1999; 99US-0129123P.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
DR P-PSDB; AAB27209.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 1; Fig 2; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase II coding
CC sequence. Its protein is thought to be involved in signal transduction in
CC the cell, and to have binding activity for other telomere-associated
CC proteins. It is possible that it plays a role in the regulation of
CC telomere length, thus affecting the replicative ability of the cell. The
CC protein is useful for ribosylating target proteins, for determining
CC tankyrase II binding activity in a sample, and for modulating telomere
CC length in a cell
XX
SQ Sequence 4493 BP; 1119 A; 685 C; 760 G; 985 T; 0 U; 944 Other;

Query Match 78.2%; Score 2969.6; DB 3; Length 4493;
Best Local Similarity 88.6%; Pred. No. 0;
Matches 3028; Conservative 1; Mismatches 363; Indels 27; Gaps 2;

```
Qy       403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db       672 AGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGCTTCAAAATGGTGCAAATNTCCAANC 731

Qy       463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db       732 ACTTTATAATGGGGGCCTTATTCCTCTTCATANTGCATGCTCTTTTGGTCATGCTAAAAN 791

Qy       523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db       792 TATCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 851

Qy       583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db       852 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 911

Qy       643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db       912 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 971

Qy       703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db       972 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1031

Qy       763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db       1032 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1091

Qy       823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db       1092 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1151

Qy       883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db       1152 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1211

Qy       943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db       1212 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1271

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db       1272 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1331

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db       1332 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1391

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db       1392 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1451

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db       1452 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1511

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db       1512 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1571

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db       1572 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1631

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db       1632 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1691

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db       1692 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1751

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db       1752 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1811

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db       1812 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1871

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db       1872 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1931

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db       1932 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1991

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db       1992 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2051

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db       2052 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2111

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db       2112 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGNGGCCTTGTACCTTTGCACAATGC 2171

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       2172 ATGTTNTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2231

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       2232 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2291
```

```
Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
             |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db       2292 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2351

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
             ||||| |||||||||||||||||||||| |||||||||||||| ||||||||||||||||
Db       2352 TACTCTTTTGGATCTTGTTAAAGATGGAGANACAGATATTCAAGATNTGCTTAGGGGAGA 2411

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
             ||||| |||| |||||||||||||||||||||||||||||||||||||||| | ||
Db       2412 TGCAGTTTTGTTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTNTNTTT 2471

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
             |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       2472 TCCTGATAATGTAAATTGCCGNGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2531

Qy       2263 AGCT------GGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGA 2316
             ||| |
Db       2532 AGGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2591

Qy       2317 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCA 2376

Db       2592 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2651

Qy       2377 TGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATG 2436

Db       2652 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2711

Qy       2437 GGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTT 2496

Db       2712 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2771

Qy       2497 GCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTT 2556

Db       2772 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2831

Qy       2557 AGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCC 2616
                                                 |||||||||||||||||||| ||||||
Db       2832 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAGCAGCCATGCCCCCATCTGTTCTGCC 2891

Qy       2617 CTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGC 2676
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2892 CTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGC 2951

Qy       2677 TCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTT 2736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2952 TCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTT 3011

Qy       2737 ATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTC 2796
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3012 ATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTC 3071

Qy       2797 CAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 2856
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3072 CAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 3131

Qy       2857 TCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATT 2916
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       3132 TCTTGGACTTGAGCACCTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGTATT 3191

Qy       2917 AGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCA 2976
             ||||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db       3192 AGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGWATCAATGCTTATGGACATAGGCA 3251

Qy       2977 CAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTT 3036
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3252 CAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTT 3311

Qy       3037 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3096
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3312 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3371

Qy       3097 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3156
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3372 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3431

Qy       3157 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 3216
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3432 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACANNNN 3491

Qy       3217 ACTATG--------------------GGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3255
                                        ||||||||||||||||||||||||||||||||||
Db       3492 NNNNNNNNGAGCCAAGATTCGGCACGAGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3551

Qy       3256 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3315
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3552 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3611

Qy       3316 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3375
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3612 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3671

Qy       3376 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3435
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3672 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3731

Qy       3436 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3495
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3732 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3791

Qy       3496 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3555
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3792 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3851

Qy       3556 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3615
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3852 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3911

Qy       3616 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3675
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3912 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3971

Qy       3676 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3735
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3972 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 4031

Qy       3736 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4032 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4090
```

```
RESULT 17
AAC66824
ID   AAC66824 standard; cDNA; 4296 BP.
XX
AC   AAC66824;
XX
DT   27-FEB-2001  (first entry)
XX
DE   Human tankyrase II coding sequence SEQ ID NO: 3.
XX
KW   Human; tankyrase II; telomere length; signal transduction; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3. .4373
FT                   /*tag= a
FT                   /transl_except= (pos:182. .184,aa:Xaa)
FT                   /transl_except= (pos:350. .352,aa:Xaa)
FT                   /transl_except= (pos:385. .387,aa:Xaa)
FT                   /transl_except= (pos:397. .399,aa:Xaa)
FT                   /transl_except= (pos:434. .436,aa:Xaa)
FT                   /transl_except= (pos:473. .475,aa:Xaa)
FT                   /transl_except= (pos:554. .556,aa:Xaa)
FT                   /transl_except= (pos:575. .580,aa:XaaXaa)
```

```
FT                   /transl_except= (pos:590. .592,aa:Xaa)
FT                   /transl_except= (pos:611. .613,aa:Xaa)
FT                   /transl_except= (pos:620. .622,aa:Xaa)
FT                   /transl_except= (pos:662. .664,aa:Xaa)
FT                   /transl_except= (pos:668. .673,aa:XaaXaa)
FT                   /transl_except= (pos:677. .679,aa:Xaa)
FT                   /transl_except= (pos:689. .691,aa:Xaa)
FT                   /transl_except= (pos:95. .697,aa:Xaa)
FT                   /transl_except= (pos:701. .703,aa:Xaa)
FT                   /transl_except= (pos:737. .742,aa:XaaXaa)
FT                   /transl_except= (pos:767. .772,aa:XaaXaa)
FT                   /transl_except= (pos:2057. .2059,aa:Xaa)
FT                   /transl_except= (pos:2231. .2233,aa:Xaa)
FT                   /transl_except= (pos:2237. .2239,aa:Xaa)
FT                   /transl_except= (pos:2261. .2263,aa:Xaa)
FT                   /transl_except= (pos:2279. .2281,aa:Xaa)
FT                   /transl_except= (pos:2345. .2350,aa:XaaXaa)
FT                   /transl_except= (pos:3035. .3037,aa:Xaa)
FT                   /transl_except= (pos:3860. .3862,aa:Xaa)
FT                   /transl_except= (pos:3881. .3883,aa:Xaa)
FT                   /transl_except= (pos:3890. .3892,aa:Xaa)
FT                   /transl_except= (pos:3896. .3898,aa:Xaa)
FT                   /transl_except= (pos:3941. .3943,aa:Xaa)
FT                   /transl_except= (pos:4001. .4003,aa:Xaa)
FT                   /transl_except= (pos:4025. .30,aa:XaaXaa)
FT                   /transl_except= (pos:4118. .4120,aa:Xaa)
FT                   /transl_except= (pos:4133. .35,aa:Xaa)
FT                   /transl_except= (pos:4172. .4174,aa:Xaa)
FT                   /transl_except= (pos:4190. .4192,aa:Xaa)
FT                   /transl_except= (pos:4199. .4201,aa:Xaa)
FT                   /transl_except= (pos:4238. .4240,aa:Xaa)
FT                   /transl_except= (pos:4256. .4258,aa:Xaa)
FT                   /transl_except= (pos:4268. .4270,aa:Xaa)
FT                   /transl_except= (pos:4283. .4285,aa:Xaa)
FT                   /transl_except= (pos:4316. .4318,aa:Xaa)
FT                   /note= "Xaa=unknown. The sequence contains a number of in
FT                   -frame termination codons, given as Xaa in the protein
FT                   sequence"
FT                   /product= "tankyrase II"
FT                   /partial
XX
PN   WO200061813-A1.
XX
PD   19-OCT-2000.
XX
PF   10-APR-2000; 2000WO-US009558.
XX
PR   09-APR-1999;   99US-0128577P.
PR   13-APR-1999;   99US-0129123P.
XX
PA   (GERO-) GERON CORP.
XX
PI   Morin GB,  Funk WD,  Piatyszek MA;
XX
DR   WPI; 2000-679503/66.
DR   P-PSDB; AAB27210.
XX
PT   Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT   the polypeptide useful for modulating or maintaining telomere length,
PT   replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS   Disclosure; Fig 3; 52pp; English.
XX
CC   The present sequence is a version of the human tankyrase II coding
CC   sequence. Its protein is thought to be involved in signal transduction in
CC   the cell, and to have binding activity for other telomere-associated
CC   proteins. It is possible that it plays a role in the regulation of
CC   telomere length, thus affecting the replicative ability of the cell. The
CC   protein is useful for ribosylating target proteins, for determining
CC   tankyrase II binding activity in a sample, and for modulating telomere
CC   length in a cell
XX
SQ   Sequence 4296 BP; 1161 A; 820 C; 925 G; 1033 T; 0 U; 357 Other;
```

```
  Query Match             78.2%;  Score 2968.2;  DB 3;  Length 4296;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches 3020;  Conservative    7;  Mismatches  381;  Indels   12;  Gaps    2;

Qy        384 CCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGA 443
              |||    ||  | :|    | |||||||||||||||||| || ||:|||||||||||| |
Db        475 CCCCGCTGCACTTYCCCGCANGTTTTGGGCGGAAAGACTTANTTRAATATTTGCTTCANA 534

Qy        444 ATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCT 503
              |||||||| |||:||| |||||||| ||||||||||||||||||||||||||||||||||
Db        535 ATGGTGCAAATGTYCAANCACGTGATNATGGGGGCCTTATTCCTCTTCATAATGCATGCT 594

Qy        504 CTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTC 563
              |||||||||:||||:||    || |||||||||| |||| ||||| |:||||||||||
Db        595 CTTTTGGTCMTGCTRAAANTATCNATCTCCTTTTGCNACATNGTGCANAMCCCAATGCTC 654

Qy        564 GAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTT 623
              |||||||||||||||||||||||| | |||||||||||||||||||||||||    ||
Db        655 GAGATAATTGGAATTATACTCCTCNCNATGAAGCTGCAATTAAAGGAAAGATTGANNNTT 714

Qy        624 GCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAG 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 GCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAG 774

Qy        684 CATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATG 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 CATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATG 834

Qy        744 AACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCAT 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 AACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCAT 894

Qy        804 TAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAG 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 TAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAG 954

Qy        864 GATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTA 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        955 GATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTA 1014

Qy        924 AAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAA 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1015 AAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAA 1074

Qy        984 CTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1075 CTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTC 1134

Qy       1044 CTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATG 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1135 CTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATG 1194

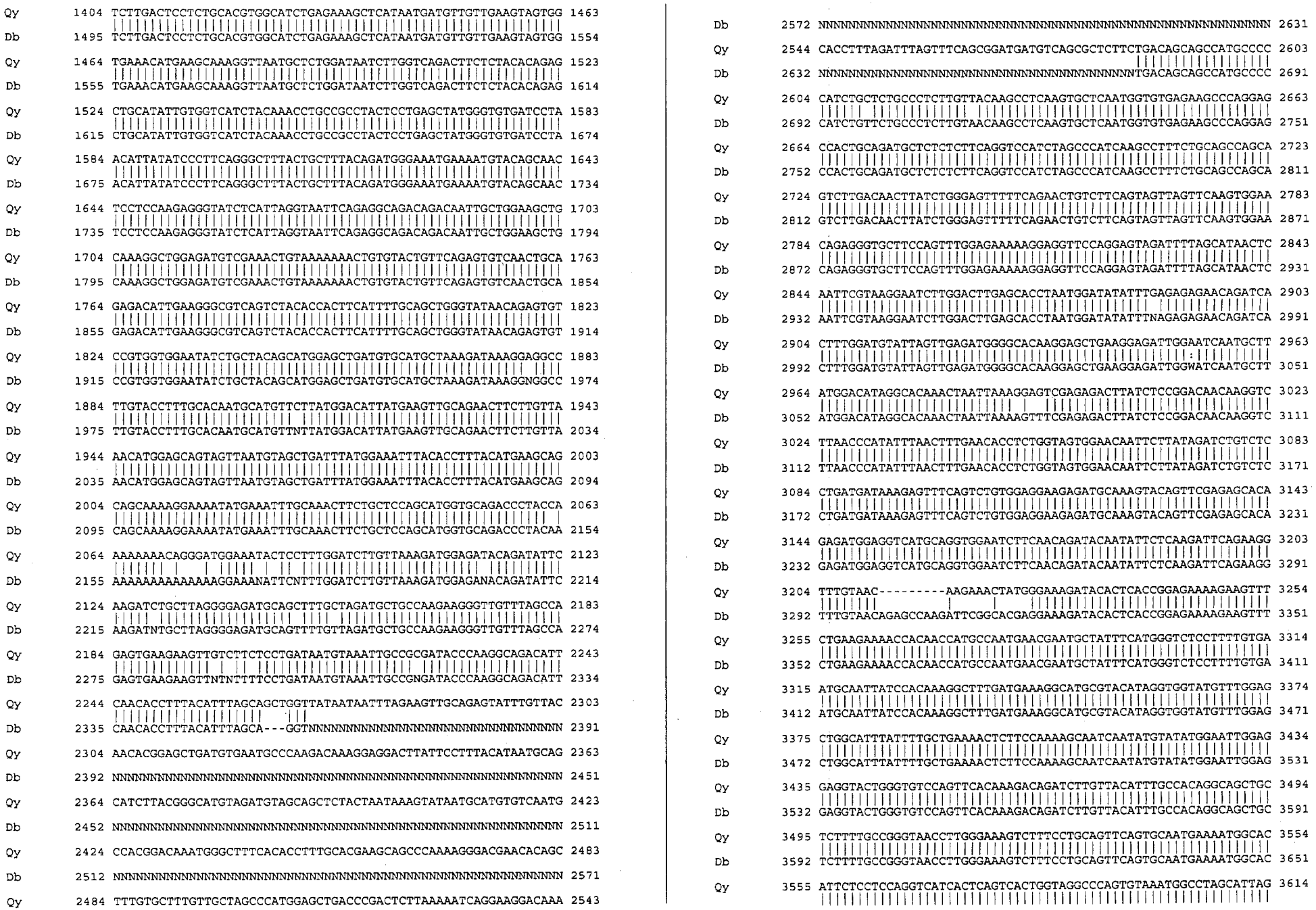
Qy       1104 GTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACAC 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1195 GTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACAC 1254

Qy       1164 CACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAC 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1255 CACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAC 1314

Qy       1224 GAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGC 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1315 GAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGC 1374

Qy       1284 ATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAA 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1375 ATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAA 1434

Qy       1344 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1435 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1494
```



```
Db       3652 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3711

Qy       3615 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3674
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3712 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3771

Qy       3675 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3734
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3772 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3831

Qy       3735 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3832 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3891
```

RESULT 18
AAC89706
ID AAC89706 standard; cDNA; 5005 BP.
XX
AC AAC89706;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human adult T-cell leukaemia cDNA MO-BC-203.
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR P-PSDB; AAB48574.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 8; Page 82-84; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 5005 BP; 1492 A; 917 C; 1015 G; 1571 T; 0 U; 10 Other;

Query Match 73.1%; Score 2774.6; DB 4; Length 5005;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2810; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

```
Qy        978 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1037
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 60

Qy       1038 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1097
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 120

Qy       1098 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1157
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 180

Qy       1158 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240

Qy       1218 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 300

Qy       1278 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 360

Qy       1338 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1397
              |||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db        361 AAAGAAAGCAAATATGTGAACTGTTGCTAAAAAAAGGATCAAACATCAATGAAAAGACTA 420

Qy       1398 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480

Qy       1458 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540

Qy       1518 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 600

Qy       1578 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 660

Qy       1638 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 720

Qy       1698 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 780

Qy       1758 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 840

Qy       1818 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 900

Qy       1878 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1937
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 960

Qy       1938 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1997
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1020

Qy       1998 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2057
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 1080

Qy       2058 CTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2117
```

```
Db       1081 CTACAAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 1140
Qy       2118 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2177
Db       1141 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 1200
Qy       2178 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2237
Db       1201 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 1260
Qy       2238 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 2297
Db       1261 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 1320
Qy       2298 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 2357
Db       1321 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 1380
Qy       2358 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 2417
Db       1381 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 1440
Qy       2418 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 2477
Db       1441 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 1500
Qy       2478 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 2537
Db       1501 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 1560
Qy       2538 GACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCA 2597
Db       1561 GACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCA 1620
Qy       2598 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 2657
Db       1621 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 1680
Qy       2658 CAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 2717
Db       1681 CAGGAGCCACTGCAGATGCTCTCT-TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 1739
Qy       2718 CCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 2777
Db       1740 CCAGCAGTCTTGACAAC-TATCTGGGAG-TTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 1797
Qy       2778 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 2837
Db       1798 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 1857
Qy       2838 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 2897
Db       1858 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 1917
Qy       2898 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 2957
Db       1918 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 1977
Qy       2958 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 3017
Db       1978 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 2037
Qy       3018 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 3077
Db       2038 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 2097
Qy       3078 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 3137
Db       2098 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 2157
Qy       3138 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 3197
Db       2158 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 2217
Qy       3198 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 3257
Db       2218 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 2277
Qy       3258 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 3317
Db       2278 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 2337
Qy       3318 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3377
Db       2338 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 2397
Qy       3378 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 3437
Db       2398 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 2457
Qy       3438 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 3497
Db       2458 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 2517
Qy       3498 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 3557
Db       2518 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 2577
Qy       3558 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 3617
Db       2578 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 2637
Qy       3618 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 3677
Db       2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 2697
Qy       3678 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 3737
Db       2698 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 2757
Qy       3738 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       2758 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2814
```

RESULT 19
AAF63919
ID AAF63919 standard; DNA; 2971 BP.
XX
AC AAF63919;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone consensus SEQ ID NO: 92.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by

PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 152-153; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 2971 BP; 915 A; 555 C; 642 G; 859 T; 0 U; 0 Other;

Query Match 64.3%; Score 2439.8; DB 5; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1352 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1412 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1472 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1532 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1592 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1652 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1712 GGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1771
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1772 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1832 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1892 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1952 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2012 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2071
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720

Qy       2072 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2132 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2192 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2252 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2372 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2432 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2491
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy       2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy       2552 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2611
             ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2612 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy       2672 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy       2732 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy       2792 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy       2852 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy       2912 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2972 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy       3032 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy       3092 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy       3152 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860

Qy       3212 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy       3272 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy       3332 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy       3392 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy       3452 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy       3512 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3572 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3692 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy       3752 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
              ||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 20
AAF63917
ID AAF63917 standard; DNA; 3353 BP.
XX
AC AAF63917;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone coding sequence SEQ ID NO: 88.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66285.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 141-144; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 3353 BP; 1034 A; 616 C; 709 G; 994 T; 0 U; 0 Other;

Query Match 64.3%; Score 2439.8; DB 5; Length 3353;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1352 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1412 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1472 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1532 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1592 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1652 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1712 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1772 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1832 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1892 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1952 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2012 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2071
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAAC 720

Qy       2072 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2132 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840
```

```
Qy      2192 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2251
Db       841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy      2252 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2311
Db       901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy      2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2371
Db       961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy      2372 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2431
Db      1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy      2432 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2491
Db      1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy      2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2551
Db      1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy      2552 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2611
Db      1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy      2612 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2671
Db      1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy      2672 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2731
Db      1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy      2732 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2791
Db      1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy      2792 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2851
Db      1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy      2852 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2911
Db      1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy      2912 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2971
Db      1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy      2972 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3031
Db      1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy      3032 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3091
Db      1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy      3092 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3151
Db      1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy      3152 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3211
Db      1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860

Qy      3212 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3271
Db      1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy      3272 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3331
Db      1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy      3332 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3391
Db      1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy      3392 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451
Db      2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy      3452 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3511
Db      2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy      3512 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3571
Db      2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy      3572 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3631
Db      2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy      3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3691
Db      2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy      3692 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3751
Db      2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy      3752 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db      2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 21
AAF63918
ID AAF63918 standard; DNA; 3799 BP.
XX
AC AAF63918;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone coding sequence SEQ ID NO: 90.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66286.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 146-150; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the

CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 3799 BP; 1157 A; 694 C; 806 G; 1142 T; 0 U; 0 Other;

Query Match 62.1%; Score 2357.8; DB 5; Length 3799;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy    1434 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 1493
Db       1 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 60

Qy    1494 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 1553
Db      61 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 120

Qy    1554 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 1613
Db     121 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 180

Qy    1614 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 1673
Db     181 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 240

Qy    1674 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1733
Db     241 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 300

Qy    1734 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 1793
Db     301 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 360

Qy    1794 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 1853
Db     361 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 420

Qy    1854 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 1913
Db     421 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 480

Qy    1914 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 1973
Db     481 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 540

Qy    1974 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 2033
Db     541 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 600

Qy    2034 TTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGG 2093
Db     601 TTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGG 660

Qy    2094 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 2153
Db     661 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 720

Qy    2154 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 2213
Db     721 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 780

Qy    2214 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 2273
Db     781 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 840

Qy    2274 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2333
Db     841 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 900

Qy    2334 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2393
Db     901 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 960

Qy    2394 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2453
Db     961 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 1020

Qy    2454 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2513
Db    1021 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 1080

Qy    2514 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2573
Db    1081 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 1140

Qy    2574 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2633
Db    1141 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 1200

Qy    2634 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2693
Db    1201 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 1260

Qy    2694 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2753
Db    1261 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 1320

Qy    2754 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2813
Db    1321 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 1380

Qy    2814 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2873
Db    1381 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 1440

Qy    2874 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2933
Db    1441 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 1500

Qy    2934 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 2993
Db    1501 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 1560

Qy    2994 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 3053
Db    1561 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 1620

Qy    3054 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 3113
Db    1621 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 1680

Qy    3114 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 3173
Db    1681 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 1740

Qy    3174 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 3233
Db    1741 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 1800

Qy    3234 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 3293
Db    1801 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 1860

Qy    3294 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 3353
Db    1861 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 1920

Qy    3354 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 3413
Db    1921 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 1980
```

```
Qy       3414 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 3473
Db       1981 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 2040

Qy       3474 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 3533
Db       2041 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 2100

Qy       3534 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 3593
Db       2101 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 2160

Qy       3594 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 3653
Db       2161 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 2220

Qy       3654 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 3713
Db       2221 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 2280

Qy       3714 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 3773
Db       2281 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 2340

Qy       3774 TACTCCTTTGCTGAAAAAAAA 3794
Db       2341 TACTCCTTTGCTGAAAAAAAA 2361
```

RESULT 22
AAF63838
ID AAF63838 standard; DNA; 3984 BP.
XX
AC AAF63838;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase1 coding sequence SEQ ID NO: 3.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66279.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 113-118; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders

XX
SQ Sequence 3984 BP; 1066 A; 1013 C; 1001 G; 904 T; 0 U; 0 Other;

Query Match 47.8%; Score 1813.4; DB 5; Length 3984;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        652 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 711

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        712 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 771

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        772 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 831

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        832 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 891

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        892 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 951

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        952 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1011

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1012 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1071

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1072 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1131

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1132 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1191

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db       1192 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1251

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db       1252 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1311

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db       1312 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1371

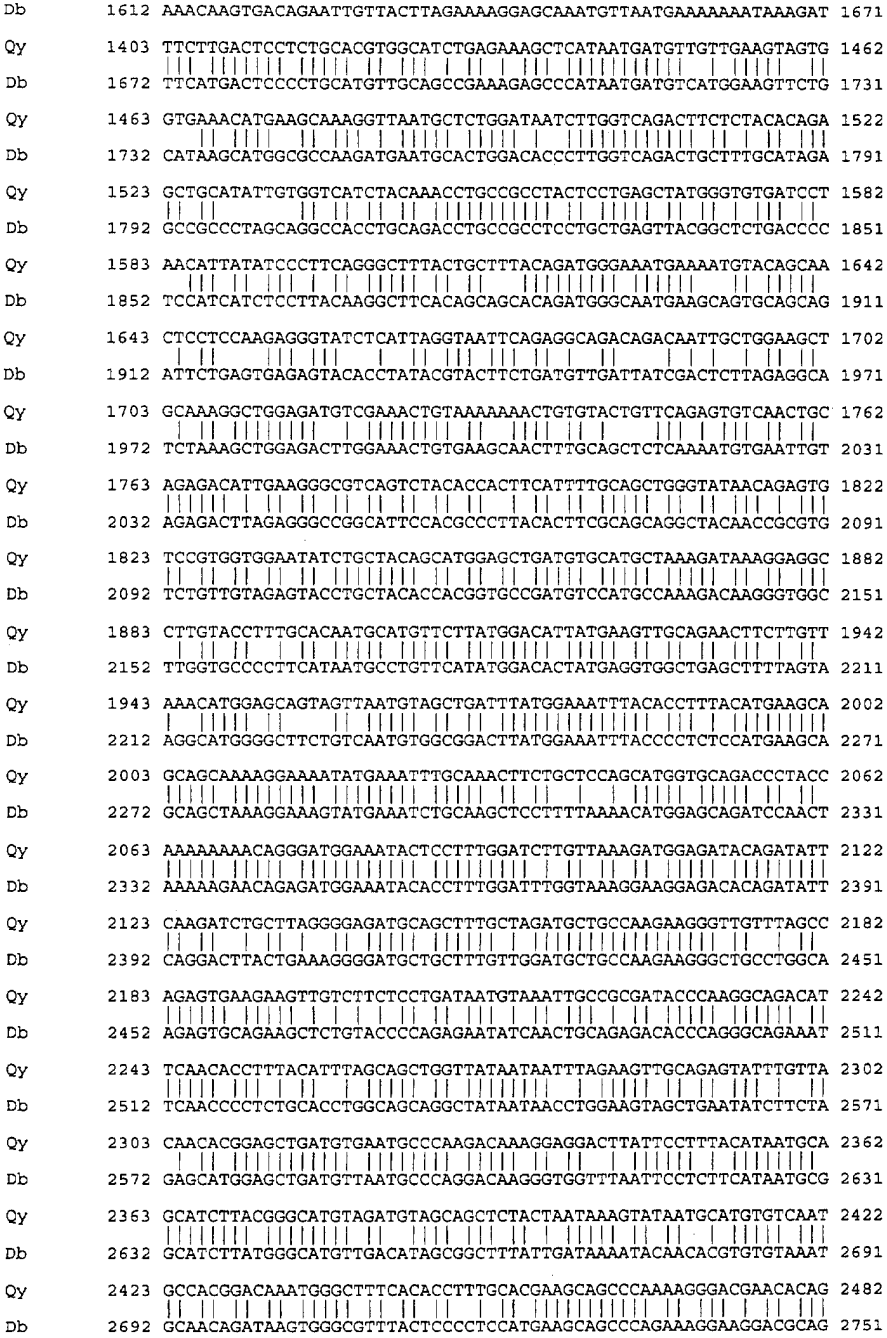
Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db       1372 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1431

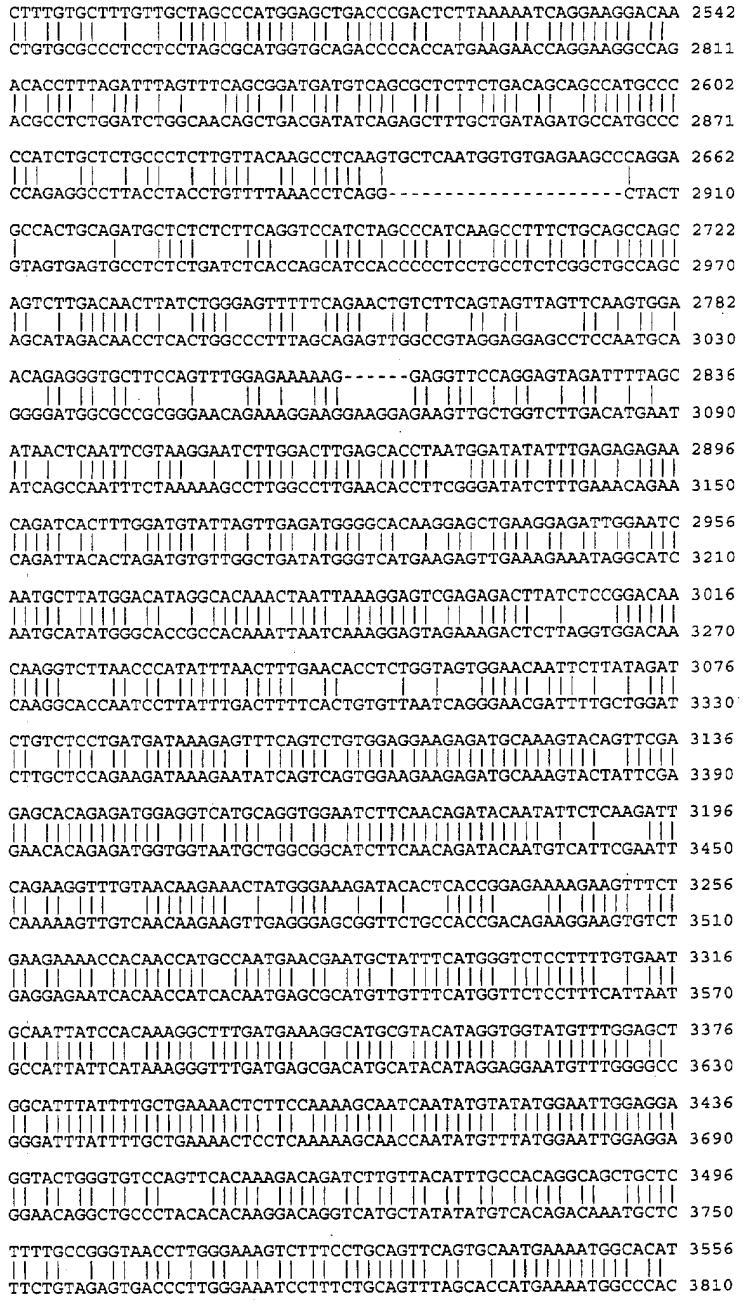
Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db       1432 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1491

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db       1492 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1551

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db       1552 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1611

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
```





```
Qy       3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616
Db       3811 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3870

Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676
Db       3871 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3930

Qy       3677 ATGAGGCCTGAAG 3689
Db       3931 ATGAAGCCAGAAG 3943
```

RESULT 23
AAZ29627
ID AAZ29627 standard; cDNA; 4134 BP.
XX
AC AAZ29627;
XX
DT 22-MAR-2000 (first entry)
XX
DE Human tankyrase cDNA.
XX
KW Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW human testis library; recombinant tankyrase; telomere length homeostasis;
KW skin atrophy; macular degeneration; atherosclerosis; gene therapy;
KW ataxia telangiectasia; drug screening; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 6. .3989
FT /*tag= a
FT /product= "Human tankyrase protein"
XX
PN WO9964606-A1.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-US012968.
XX
PR 10-JUN-1998; 98US-00095225.
PR 17-AUG-1998; 98US-00135233.
PR 19-NOV-1998; 98US-00196387.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2000-116549/10.
DR P-PSDB; AAY44402.
XX
PT New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT length, used to identify modulators, e.g. for inhibiting growth of
PT cancer.
XX
PS Claim 5; Page 111-112; 133pp; English.
XX
CC The present cDNA sequence encodes human tankyrase, a TRF1 binding
CC protein. This cDNA was isolated from a human testis library. Tankyrase is
CC involved in regulation of telomere length through TRF1 and may directly
CC modulate the effect of TRF1. This gene is used to produce recombinant
CC tankyrase, or its fragments, and these are used to identify specific
CC modulators which are potential drugs to control telomere length
CC homeostasis for countering telomere shortening associated with aging
CC (e.g. atrophy of skin, macular degeneration or atherosclerosis) or
CC diseases such as ataxia telangiectasia. This can also be used to identify
CC homologous genes and in gene therapy (including antisense or ribozyme
CC expression). Transgenic animals having at least one disrupted endogenous
CC allele for tankyrase are used in drug screening
XX
SQ Sequence 4134 BP; 1112 A; 1041 C; 1034 G; 947 T; 0 U; 0 Other;

Query Match 47.8%; Score 1813.4; DB 3; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

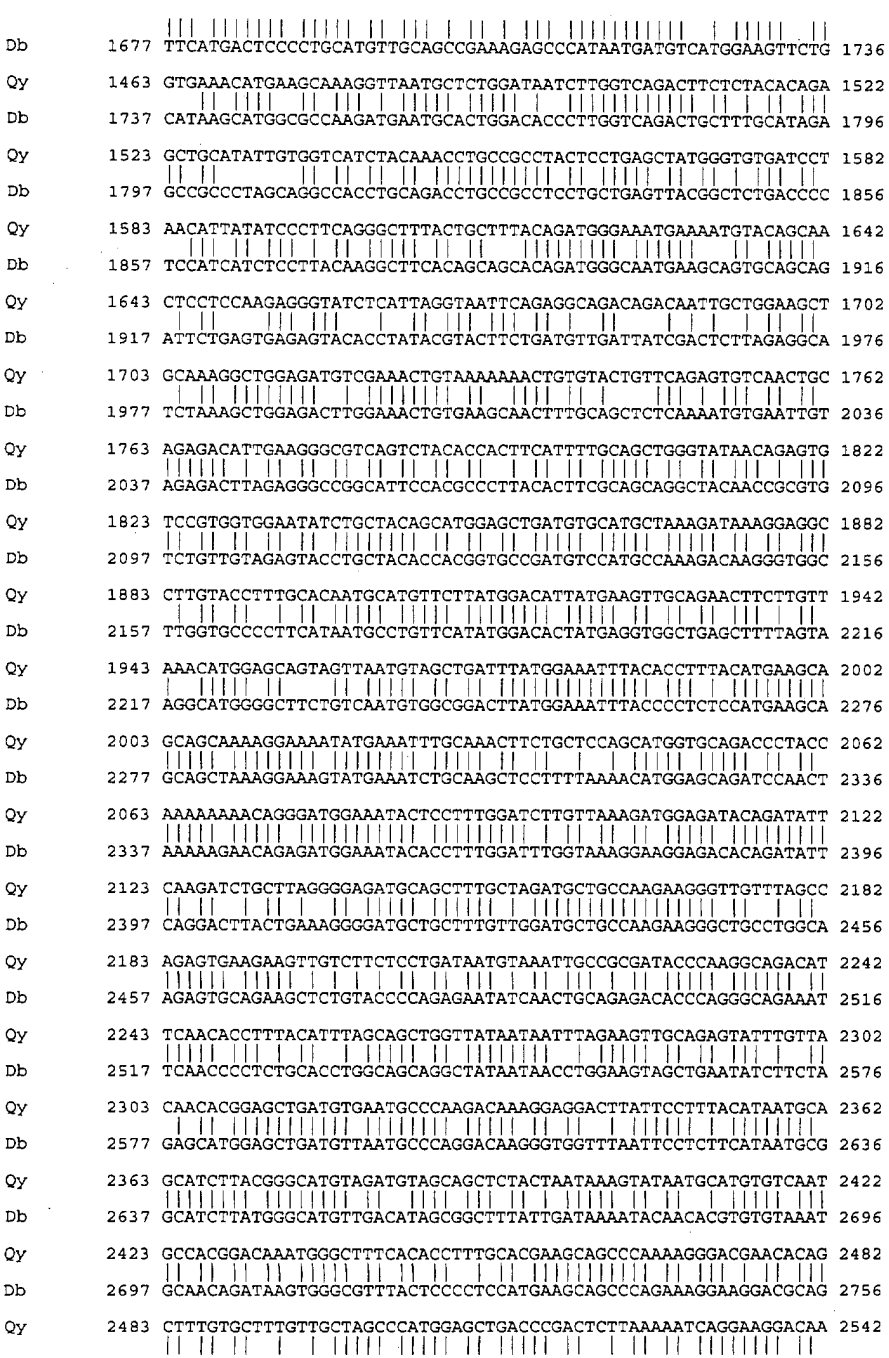
Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
```


Db 1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736
Qy 1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db 1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796
Qy 1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
Db 1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856
Qy 1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
Db 1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916
Qy 1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
Db 1917 ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCA 1976
Qy 1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
Db 1977 TCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGT 2036
Qy 1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
Db 2037 AGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTG 2096
Qy 1823 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1882
Db 2097 TCTGTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAGACAAGGGTGGC 2156
Qy 1883 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1942
Db 2157 TTGGTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTGAGCTTTTAGTA 2216
Qy 1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2002
Db 2217 AGGCATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTCTCCATGAAGCA 2276
Qy 2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2062
Db 2277 GCAGCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAGCAGATCCAACT 2336
Qy 2063 AAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATT 2122
Db 2337 AAAAAGAACAGAGATGGAAATACACCTTTGGATTTGGTAAAGGAAGGAGACACAGATATT 2396
Qy 2123 CAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCC 2182
Db 2397 CAGGACTTACTGAAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGGGCTGCCTGGCA 2456
Qy 2183 AGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACAT 2242
Db 2457 AGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAAT 2516
Qy 2243 TCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTA 2302
Db 2517 TCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTTCTA 2576
Qy 2303 CAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCA 2362
Db 2577 GAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCTCTTCATAATGCG 2636
Qy 2363 GCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAAT 2422
Db 2637 GCATCTTATGGGCATGTTGACATAGCGGCTTTATTGATAAAATACAACACGTGTGTAAAT 2696
Qy 2423 GCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAG 2482
Db 2697 GCAACAGATAAGTGGGCGTTTACTCCCCTCCATGAAGCAGCCCAGAAAGGAAGGACGCAG 2756
Qy 2483 CTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAA 2542
Db 2757 CTGTGCGCCCTCCTCCTAGCGCATGGTGCAGACCCCACCATGAAGAACCAGGAAGGCCAG 2816
Qy 2543 ACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCC 2602
Db 2817 ACGCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCCC 2876
Qy 2603 CCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGA 2662
Db 2877 CCAGAGGCCTTACCTACCTGTTTTAAACCTCAGG----------CTACT 2915
Qy 2663 GCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGC 2722
Db 2916 GTAGTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGC 2975
Qy 2723 AGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGA 2782
Db 2976 AGCATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCA 3035
Qy 2783 ACAGAGGGTGCTTCCAGTTTGGAGAAAAAG------GAGGTTCCAGGAGTAGATTTTAGC 2836
Db 3036 GGGGATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAAT 3095
Qy 2837 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 2896
Db 3096 ATCAGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAA 3155
Qy 2897 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 2956
Db 3156 CAGATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATC 3215
Qy 2957 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 3016
Db 3216 AATGCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAA 3275
Qy 3017 CAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 3076
Db 3276 CAAGGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGAT 3335
Qy 3077 CTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGA 3136
Db 3336 CTTGCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGA 3395
Qy 3137 GAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATT 3196
Db 3396 GAACACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATT 3455
Qy 3197 CAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCT 3256
Db 3456 CAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCT 3515
Qy 3257 GAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAAT 3316
Db 3516 GAGGAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAAT 3575
Qy 3317 GCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCT 3376
Db 3576 GCCATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCC 3635
Qy 3377 GGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA 3436
Db 3636 GGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGA 3695
Qy 3437 GGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTC 3496
Db 3696 GGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTC 3755
Qy 3497 TTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACAT 3556
Db 3756 TTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCAC 3815
Qy 3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616
Db 3816 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3875

```
Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676
             |||||||| || |||||||||||||||||| || || |||||| | || ||||||||||||
Db       3876 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3935

Qy       3677 ATGAGGCCTGAAG 3689
              |||| ||| ||||
Db       3936 ATGAAGCCAGAAG 3948
```

RESULT 24
AAC66826
ID AAC66826 standard; DNA; 4134 BP.
XX
AC AAC66826;
XX
DT 27-FEB-2001 (first entry)
XX
DE Human tankyrase I coding sequence SEQ ID NO: 7.
XX
KW Human; tankyrase II; telomere length; signal transduction; ds.
XX
OS Homo sapiens.
XX
PN WO200061813-A1.
XX
PD 19-OCT-2000.
XX
PF 10-APR-2000; 2000WO-US009558.
XX
PR 09-APR-1999; 99US-0128577P.
PR 13-APR-1999; 99US-0129123P.
XX
PA (GERO-) GERON CORP.
XX
PI Morin GB, Funk WD, Piatyszek MA;
XX
DR WPI; 2000-679503/66.
XX
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide encoding
PT the polypeptide useful for modulating or maintaining telomere length,
PT replicative capacity, apoptosis, chromosome packing or gene expression.
XX
PS Claim 1; Fig 6; 52pp; English.
XX
CC The present sequence is a version of the human tankyrase I coding
CC sequence. The invention relates to the isolation of the protein and
CC coding sequences of human tankyrase II. This protein is thought to be
CC involved in signal transduction in the cell, and to have binding activity
CC for other telomere-associated proteins. It is possible that it plays a
CC role in the regulation of telomere length, thus affecting the replicative
CC ability of the cell. The protein is useful for ribosylating target
CC proteins, for determining tankyrase II binding activity in a sample, and
CC for modulating telomere length in a cell
XX
SQ Sequence 4134 BP; 1112 A; 1041 C; 1034 G; 947 T; 0 U; 0 Other;

  Query Match   47.8%; Score 1813.4; DB 3; Length 4134;
  Best Local Similarity   72.3%; Pred. No. 0;
  Matches 2395; Conservative   0; Mismatches 891; Indels   27; Gaps   2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
              || | |  ||  |    |  |||||||||  |||| || || || ||| | || || |||
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
              |  |||||  | ||||||  || ||||||||||| || || || || ||||||||||| ||
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
              |||||||| |||||||| || || | ||| | ||| | || || || || || ||||||
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

```
Qy   563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
              | ||||| ||||| ||||| ||||| |||||||||| |||||||| ||||| |||||
Db   837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy   623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
              ||||||||||| | ||||| |||||||| |||| ||| || || || ||||| | ||
Db   897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy   683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
              || |||| | |||||||| || || ||||| || ||||| |||||||| ||||||||
Db   957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy   743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
              |||||| ||||| ||| |||||||| |||||||||||| | |||||| |||| || ||
Db   1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy   803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
              ||||||| || ||||| ||||||||||| ||||||| ||||| |||||| | |||||
Db   1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy   863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
              || || |||||||| || || |||| | || || ||||| |||||||| |||||
Db   1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy   923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
              ||||| ||||||| || || || | || |||| ||||| ||||| ||||||||||
Db   1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy   983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
              || ||||| | | |||||||| || ||||| ||||| ||||| | ||||| || |||
Db   1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy   1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
              || || || ||||| ||||| |||||| | || ||||| || ||| | | | || ||
Db   1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy   1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
              || || || || || | |||| || || |||||||| | || ||||||| ||
Db   1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy   1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
              || || | | || ||||| | |||||||||||| || || || || ||||| ||
Db   1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy   1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
              ||||||| || | || | | ||||| ||| ||||||||| | ||||||||
Db   1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy   1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
              || || || | |||||||||||| |||| ||||||| || |||| |||||||| |
Db   1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy   1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
              || ||| | ||| |||| || ||||||||||||| | |||||||| | |||||||
Db   1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

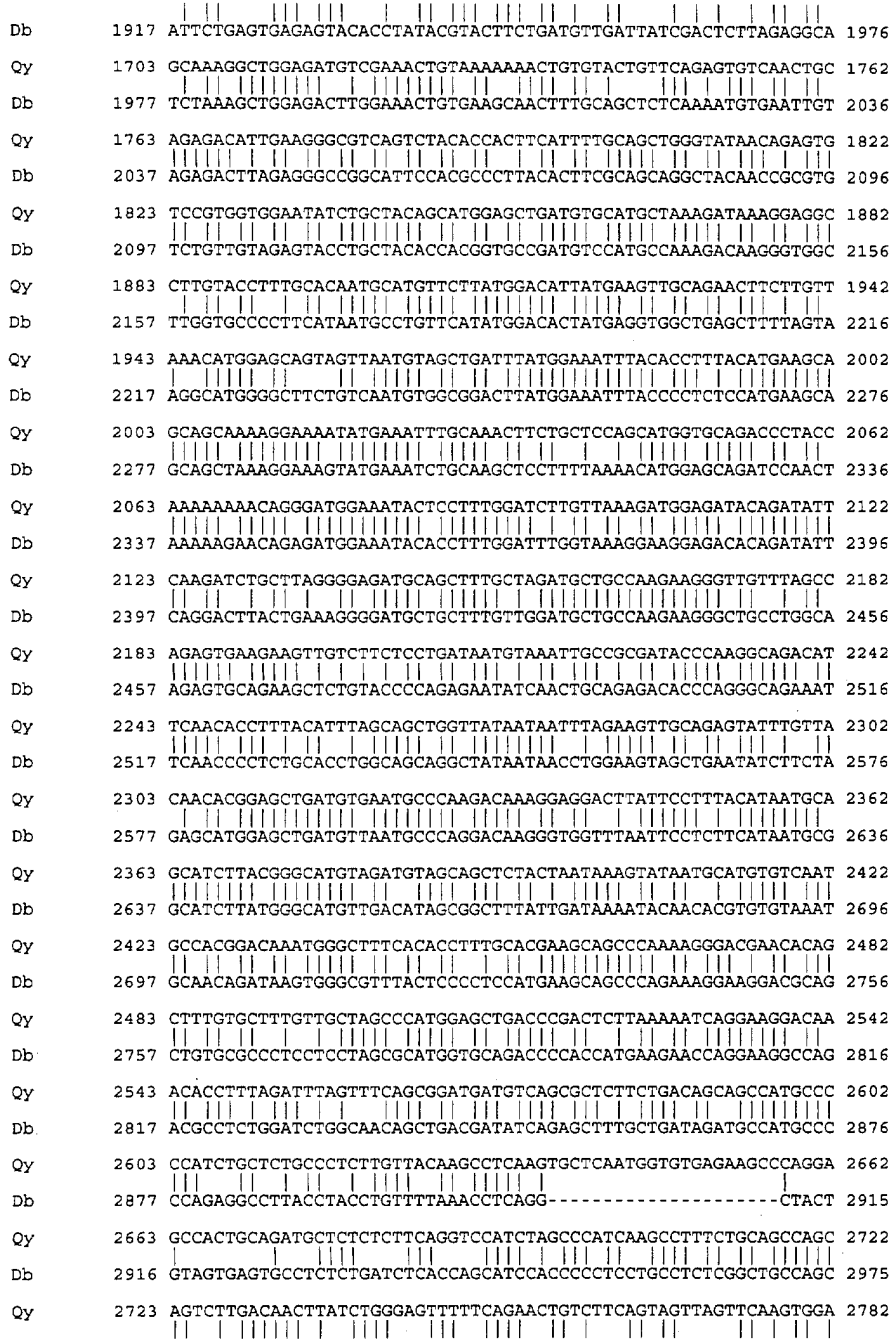
Qy   1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
              ||| ||||||| ||||| || ||| | || | ||| ||||||||||| | ||||| ||
Db   1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736

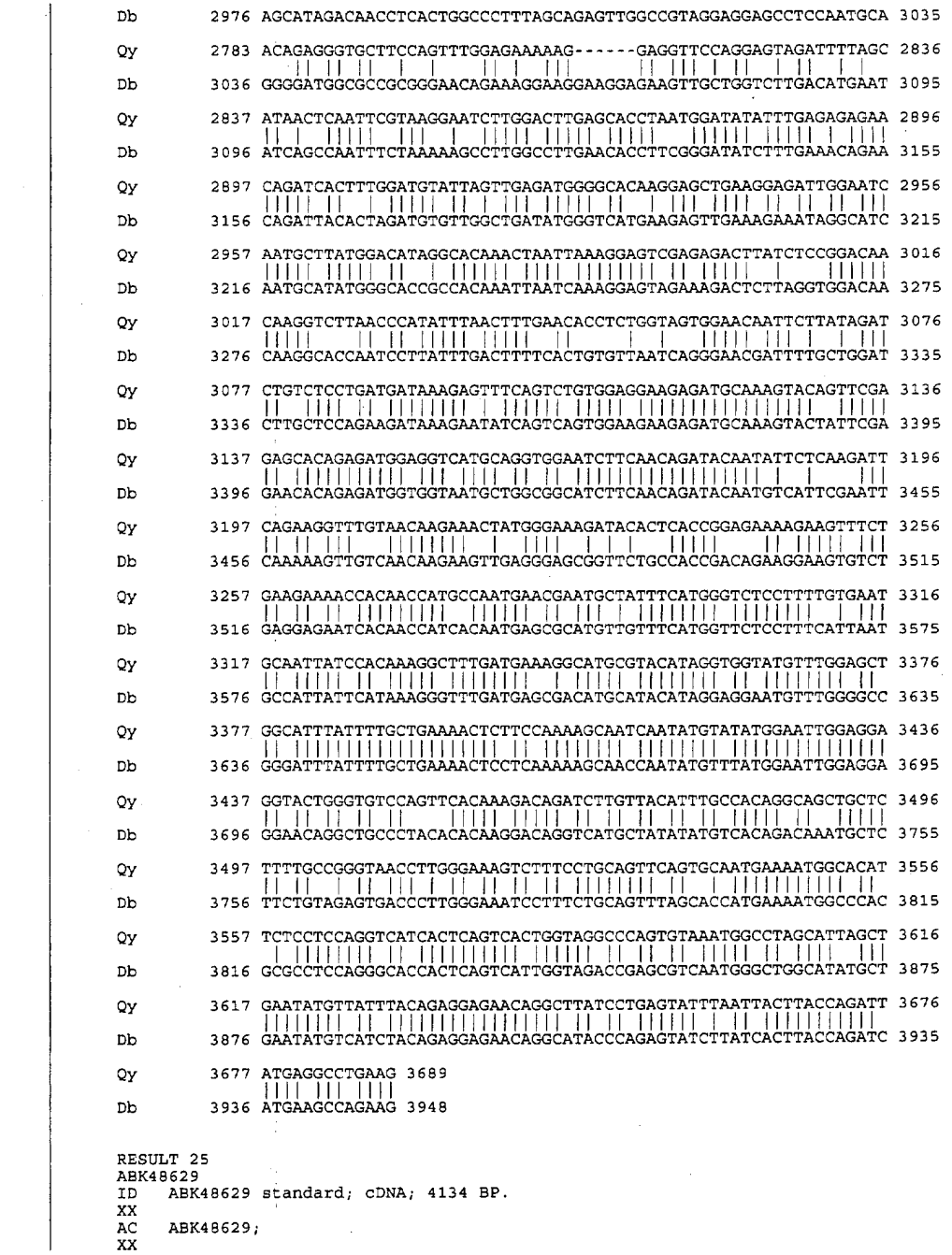
Qy   1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
              || |||| || |||| | ||||| ||||| |||||||||||| || | || |||||||
Db   1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796

Qy   1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
              || || || || || || |||||||||||| || ||||| || || | ||| ||
Db   1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856

Qy   1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
                ||| || ||| | || ||||| || || |||||||| |||||| || |||||
Db   1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916

Qy   1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
```
```

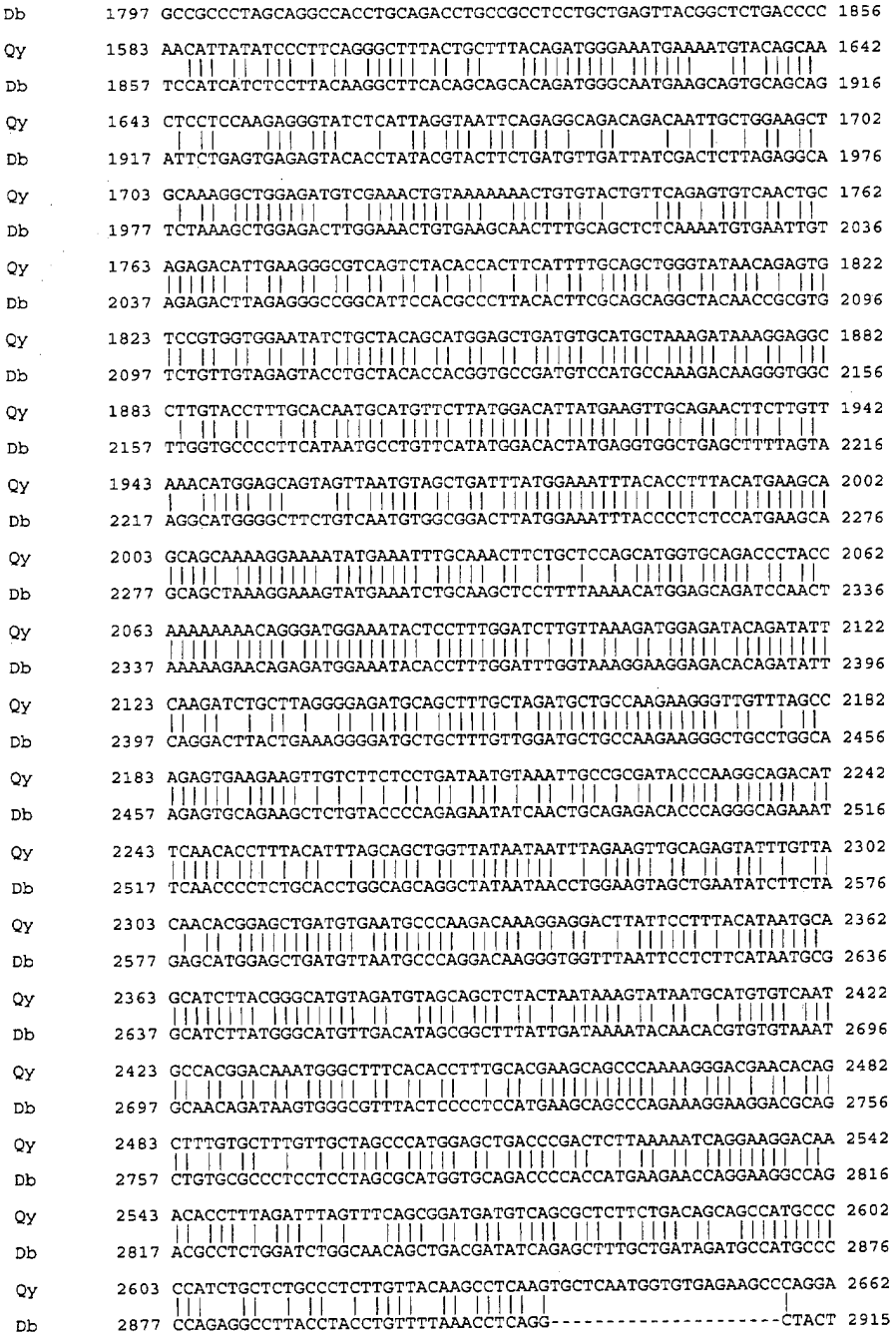


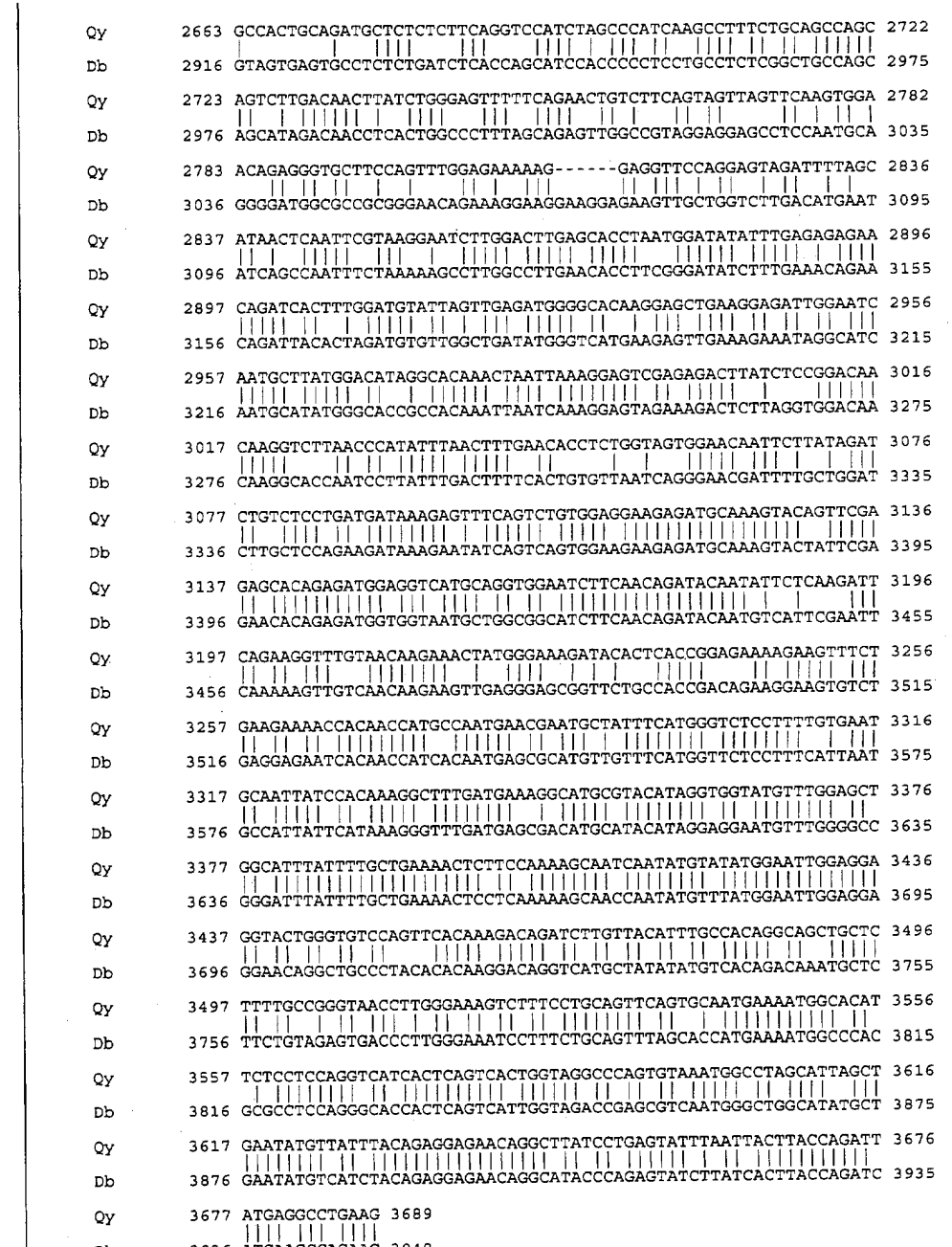


RESULT 25
ABK48629
ID ABK48629 standard; cDNA; 4134 BP.
XX
AC ABK48629;
XX

DT 12-AUG-2002 (first entry)
XX
DE Human tankyrase cDNA.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; ss; gene; cytostatic;
KW ADP ribosylation; telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 6. .3989
FT /*tag= a
FT /product= "Tankyrase"
FT 3'UTR 3990. .4134
FT /*tag= b
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR P-PSDB; AAU79537.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 127-129; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC sequence presented is the human tankyrase cDNA
XX
SQ Sequence 4134 BP; 1112 A; 1041 C; 1034 G; 947 T; 0 U; 0 Other;

Query Match 47.8%; Score 1813.4; DB 6; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;





```
RESULT 26
AAC85295
ID   AAC85295 standard; cDNA; 3254 BP.
XX
AC   AAC85295;
XX
DT   29-MAR-2001  (first entry)
XX
DE   Mouse SPANK cDNA.
XX
KW   SPANK; SAM; sterile alpha motif; PARP; insulin resistance;
KW   poly adenosine diphosphate-ribose polymerase; catalytic domain; ANK;
KW   ankyrin repeat; cytosol; insulin-responsive aminopeptidase; IRAP; GLUT4;
KW   adipocyte; insulin signalling pathway; hyperlipidaemia;
KW   glucose intolerance; atheromatous disease; atherosclerosis; obesity;
KW   cardiac insufficiency; coronary insufficiency; stroke;
KW   high blood pressure; non-insulin dependent diabetes; hypertension;
KW   hyperuricemia; Syndrome X; muscular dystrophy; muscle atrophy; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             308. .3254
FT                   /*tag= a
FT                   /product= "Murine SPANK"
FT                   /transl_except= (pos:1754,aa:Gly)
FT                   /note= "This codon has an apparent 2 nucleotide deletion,
FT                   which alters the reading frame"
FT                   /transl_except= (pos:2145,aa:DTQIGKLLIG)
FT                   /note= "The nucleotides encoding residues 614 to 623 are
FT                   absent, which alters the reading frame"
FT                   /transl_except= (pos:2602. .2606,aa:Met)
FT                   /note= "This codon has an apparent 2 nucleotide
FT                   insertion, which alters the reading frame"
XX
PN   WO200077225-A1.
XX
PD   21-DEC-2000.
XX
PF   09-JUN-2000; 2000WO-US015926.
XX
PR   11-JUN-1999;   99US-0138957P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Chi N,  Lodish HF;
XX
DR   WPI; 2001-091404/10.
DR   P-PSDB; AAB47023.
XX
PT   New insulin signaling protein SPANK, useful for reducing body mass,
PT   glucose intolerance or insulin resistance and for preventing or treating
PT   obesity-related and muscle-related diseases.
XX
PS   Claim 3; Fig 5A; 65pp; English.
XX
CC   This sequence represents the mouse SPANK cDNA. The SPANK protein
CC   comprises 3 domains: (a) a SAM (sterile alpha motif) domain; (b) a PARP
CC   (poly adenosine diphosphate-ribose polymerase) catalytic domain; and (c)
CC   an ANK domain composed of ankyrin repeats. SPANK is a cytosolic protein
CC   which can poly(ADP-ribosyl)ate itself. SPANK binds insulin-responsive
CC   aminopeptidase (IRAP) and modulates translocation of GLUT4 in the
CC   perinuclear region of adipocytes. It is an effector in the insulin
CC   signalling pathway in eukaryotic cells. SPANK is useful for reducing body
CC   mass, reducing glucose intolerance or insulin resistance, for preventing
CC   or treating obesity-related diseases or disorders, such as obesity,
CC   cardiac insufficiency, coronary insufficiency, stroke, hypertension,
CC   atheromatous disease, atherosclerosis, high blood pressure, non-insulin
CC   dependent diabetes, hyperlipidaemia, hyperuricemia and Syndrome X and is
CC   also useful for preventing or treating muscle-related diseases or
CC   disorders, such as muscular dystrophy, muscle atrophy and muscle fatigue.
CC   Antibodies immunospecific for SPANK are useful for detecting the presence
CC   of SPANK polypeptide in a biological sample
XX
SQ   Sequence 3254 BP; 871 A; 746 C; 863 G; 774 T; 0 U; 0 Other;

  Query Match            47.3%;  Score 1794.8;  DB 4;  Length 3254;
  Best Local Similarity  76.2%;  Pred. No. 0;
  Matches 2518;  Conservative    0;  Mismatches  232;  Indels  554;  Gaps    3;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||| |||| || |||||||||||| | |||||||| ||||||| ||| |||||
Db        505 AGGTTTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGCAAGC 564

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              |||||||||||||| |||||||||||||||| |||||||||||||||||||||||||||
Db        565 GCGTGATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTGAAGT 624

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||| ||||||| ||| |||||||||||| |||||||| ||||||||||||||||
Db        625 AGTCAATCTTCTTTTGCAACACGGTGCAGACCCAAATGCTCGGGATAATTGGAATTATAC 684

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        685 TCCTCTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTACAGCA 744

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        745 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGACCC 804

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              |||||||||||| || || |||||||| |||||||||||||||| | ||||||||||||
Db        805 ATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGAAAGATGAACTTCTGGAAAGTGCCAG 864

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              |||||||||||| |||||||||||||| || || |||||||| ||||||||||||| ||
Db        865 GAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCCATGC 924

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||| ||
Db        925 AAGTGATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGATACAACAGAGTAAAAAT 984

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              |||||| |||||||||| ||||| |||||||||||||||||||||||| |||||||||
Db        985 TGTACAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATCTGGT 1044

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              |||| |||||||||||||| ||||| |||||||||||||||||||||| ||||||||||
Db       1045 ACCACTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCA 1104

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              |||||||||| |||||||||||| ||||||| |||||||||||||||||||||||||||
Db       1105 CGGTGCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAGCTTC 1164

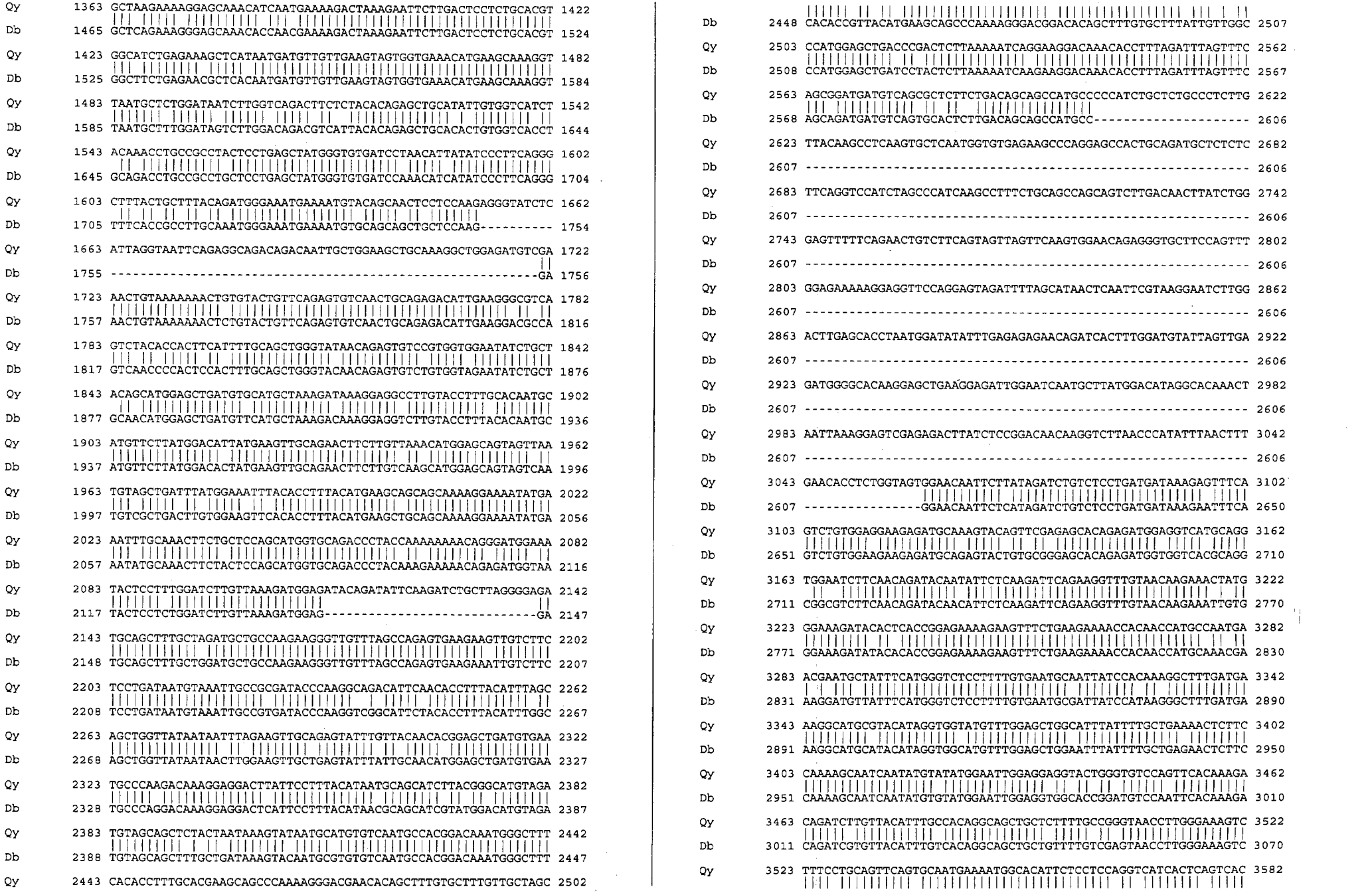
Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              |||||||||| |||||||||||||| ||| ||||||| || ||||| ||||| |||||||
Db       1165 TAAGAACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCAACCCTGCT 1224

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              || |||||||||||||||||||||||||||||| |||| ||||| ||||| | || ||||||||
Db       1225 GAACTGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAAGATT 1284

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              | | |||||||||||||||| || |||||||||||||||||||||||||||||||| |
Db       1285 ATCGTATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTACAAG 1344

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              |||||||||||||||||||||||||||||||| |||||||| |||||||||||||||||
Db       1345 AATCAAAAAACATCTTTCTCTGGAAATGGTGAATTTCAAACATCCTCAGACACATGAAAC 1404

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              |||| |||||||||||||||||||||||||||||||||||||| |||||||||||| |
Db       1405 AGCACTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAGATATGTGAACTGCT 1464
```



```
Db       3071 TTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGGTCAC 3130

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3131 TGGCCGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3190

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             |||||||||||| ||| |||||||||||||||| | ||||||||||||||||| |||||
Db       3191 GGCTTATCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAGGTATGGTTGATGG 3250

Qy       3703 ATAA 3706
                | |
Db       3251 CTGA 3254


RESULT 27
ADE53995
ID   ADE53995 standard; cDNA; 4435 BP.
XX
AC   ADE53995;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human prostate cancer cDNA #342.
XX
KW   Human; prostate cancer; ss; cDNA combination; differential expression;
KW   gene.
XX
OS   Homo sapiens.
XX
PN   US2003190640-A1.
XX
PD   09-OCT-2003.
XX
PF   29-MAY-2002; 2002US-00252157.
XX
PR   31-MAY-2001; 2001US-0295048P.
XX
PA   (FARI/) FARIS M.
PA   (PEAR/) PEARSON C I.
XX
PI   Faris M,  Pearson CI;
XX
DR   WPI; 2003-831619/77.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   prostate cancer, useful for diagnosing, treating or monitoring the
PT   progression of treatment of prostate cancer.
XX
PS   Claim 5; SEQ ID NO 342; 42pp; English.
XX
CC   The invention relates to a combination comprising a number of cDNAs
CC   expressed in prostate cancer. The invention also relates to a method for
CC   detecting differential expression of one or more cDNAs in a sample
CC   containing nucleic acids by hybridising a substrate with the nucleic
CC   acids, thus forming one or more hybridisation complexes, detecting
CC   hybridisation complex formation and comparing the complexes formed with
CC   standard complexes, where differences between the standard and the sample
CC   complex formation indicate differential expression of cDNAs in the
CC   sample. The differential expression is diagnostic of prostate cancer. The
CC   invention also relates to proteins and antibodies related to the cDNAs.
CC   The combination is useful for diagnosing, treating or monitoring the
CC   progression of treatment of prostate cancer. The antibodies are useful
CC   for detecting prostate cancer. This sequence represents a human prostate
CC   cancer cDNA of the invention.
XX
SQ   Sequence 4435 BP; 1352 A; 757 C; 851 G; 1473 T; 0 U; 2 Other;

  Query Match             40.4%;  Score 1533.8;  DB 9;  Length 4435;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 1562;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       2186 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2245
```

```
                 || || |||| |   | |  ||| |  | |          | |  |||| |   | |
Db        289 GTTAACAAGTGGAAGTATTATGACATAGGACTAAAAAAGAAAAAACAAGTAAAGTAATTT 348

Qy       2246 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2305
                 ||   |  |  ||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TACTTCTTTTCTAAATTAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 408

Qy       2306 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 468

Qy       2366 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 528

Qy       2426 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 588

Qy       2486 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 648

Qy       2546 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 708

Qy       2606 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 768

Qy       2666 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 828

Qy       2726 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 888

Qy       2786 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 948

Qy       2846 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 1008

Qy       2906 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1009 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 1068

Qy       2966 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 1128

Qy       3026 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 1188

Qy       3086 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 1248

Qy       3146 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1249 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 1308

Qy       3206 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1309 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 1368

Qy       3266 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1369 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 1428

Qy       3326 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3385
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1429 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 1488

Qy       3386 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3445
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1489 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 1548

Qy       3446 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3505
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1549 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 1608

Qy       3506 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3565
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1609 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 1668

Qy       3566 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3625
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1669 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 1728

Qy       3626 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3685
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1729 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 1788

Qy       3686 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3745
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1789 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 1848

Qy       3746 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1849 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 1897
```

```
RESULT 28
AAZ29632
ID   AAZ29632 standard; DNA; 4491 BP.
XX
AC   AAZ29632;
XX
DT   22-MAR-2000  (first entry)
XX
DE   Human truncated tankyrase-1 DNA.
XX
KW   Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW   recombinant tankyrase; telomere length homeostasis; skin atrophy;
KW   macular degeneration; atherosclerosis; gene therapy;
KW   ataxia telangiectasia; drug screening; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             6. .2027
FT                   /*tag=  a
FT                   /product= "Human truncated tankyrase"
XX
PN   WO9964606-A1.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-US012968.
XX
PR   10-JUN-1998;   98US-00095225.
PR   17-AUG-1998;   98US-00135233.
PR   19-NOV-1998;   98US-00196387.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   De Lange T,  Smith S;
XX
DR   WPI; 2000-116549/10.
DR   P-PSDB; AAY44403.
XX
PT   New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT   length, used to identify modulators, e.g. for inhibiting growth of
PT   cancer.
XX
PS   Disclosure; Page 119-120; 133pp; English.
XX
CC   The present sequence encodes a truncated human tankyrase, a TRF1 binding
CC   protein. This DNA comprises the coding sequence for amino acid residues 1
CC   -640 of human tankyrase. Tankyrase is involved in regulation of telomere
CC   length through TRF1 and may directly modulate the effect of TRF1. This
CC   gene is used to produce recombinant tankyrase, or its fragments, and
CC   these are used to identify specific modulators which are potential drugs
CC   to control telomere length homeostasis for countering telomere shortening
CC   associated with aging (e.g. atrophy of skin, macular degeneration or
CC   atherosclerosis) or diseases such as ataxia telangiectasia. This can also
CC   be used to identify homologous genes and in gene therapy (including
CC   antisense or ribozyme expression). Transgenic animals having at least one
CC   disrupted endogenous allele for tankyrase are used in drug screening
XX
SQ   Sequence 4491 BP; 1212 A; 1123 C; 1115 G; 1041 T; 0 U; 0 Other;
```

Query Match 38.2%; Score 1449.6; DB 3; Length 4491;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

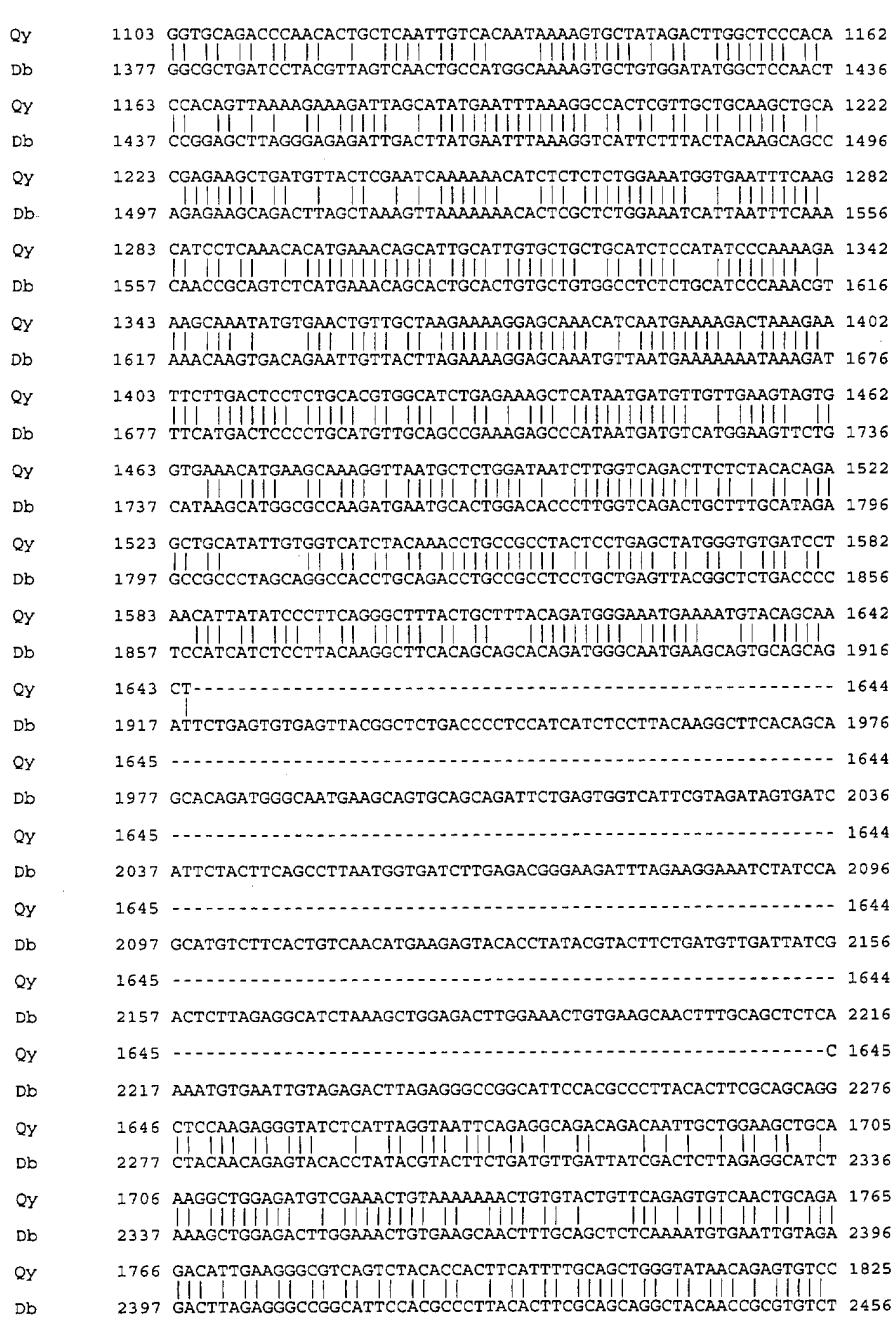
Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

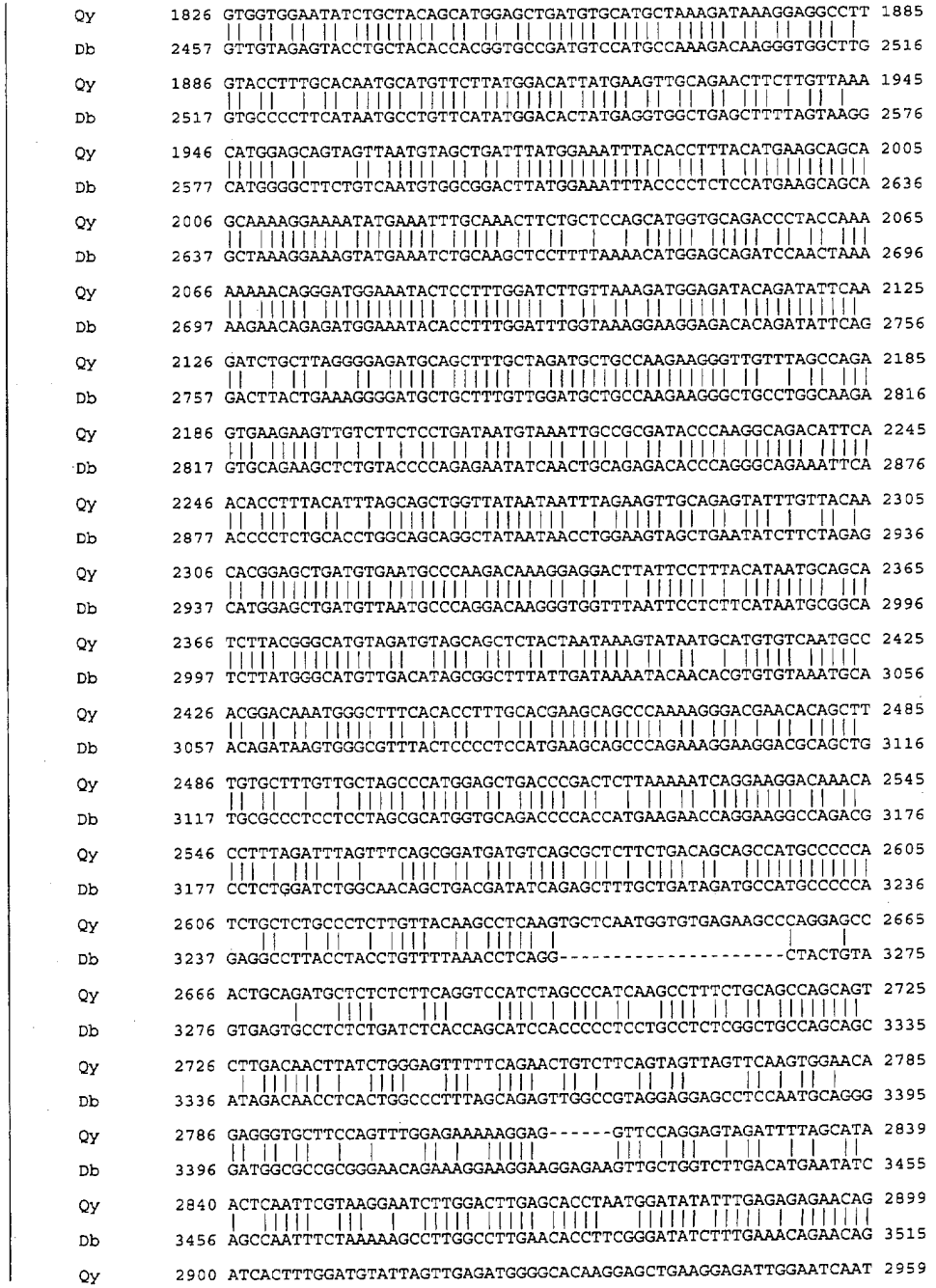
Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376
```





```
Db       3516 ATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATCAAT 3575

Qy       2960 GCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAA 3019

Db       3576 GCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAA 3635

Qy       3020 GGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTG 3079

Db       3636 GGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTT 3695

Qy       3080 TCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAG 3139

Db       3696 GCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAA 3755

Qy       3140 CACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAG 3199

Db       3756 CACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAA 3815

Qy       3200 AAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAA 3259

Db       3816 AAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAG 3875

Qy       3260 GAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCA 3319

Db       3876 GAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCC 3935

Qy       3320 ATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGC 3379

Db       3936 ATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGG 3995

Qy       3380 ATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGT 3439

Db       3996 ATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGA 4055

Qy       3440 ACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTT 3499

Db       4056 ACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTC 4115

Qy       3500 TGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCT 3559

Db       4116 TGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCG 4175

Qy       3560 CCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAA 3619

Db       4176 CCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAA 4235

Qy       3620 TATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATG 3679

Db       4236 TATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATG 4295

Qy       3680 AGGCCTGAAG 3689

Db       4296 AAGCCAGAAG 4305
```

RESULT 29
ABK48634
ID ABK48634 standard; cDNA; 4491 BP.
XX
AC ABK48634;
XX
DT 12-AUG-2002 (first entry)
XX
DE Truncated human tankyrase #1 cDNA.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; ss; gene; cytostatic;
KW ADP ribosylation; telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT CDS 6. .2027
FT /*tag= a
FT /product= "Truncated tankyrase #1"
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR P-PSDB; AAU79538.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 134-138; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC sequence presented is the truncated human tankyrase #1 cDNA. Note:
CC Nucleotides 1-1925 are the same as the wild-type human tankyrase sequence
CC (ABK48629)
XX
SQ Sequence 4491 BP; 1212 A; 1123 C; 1115 G; 1041 T; 0 U; 0 Other;

Query Match 38.2%; Score 1449.6; DB 6; Length 4491;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

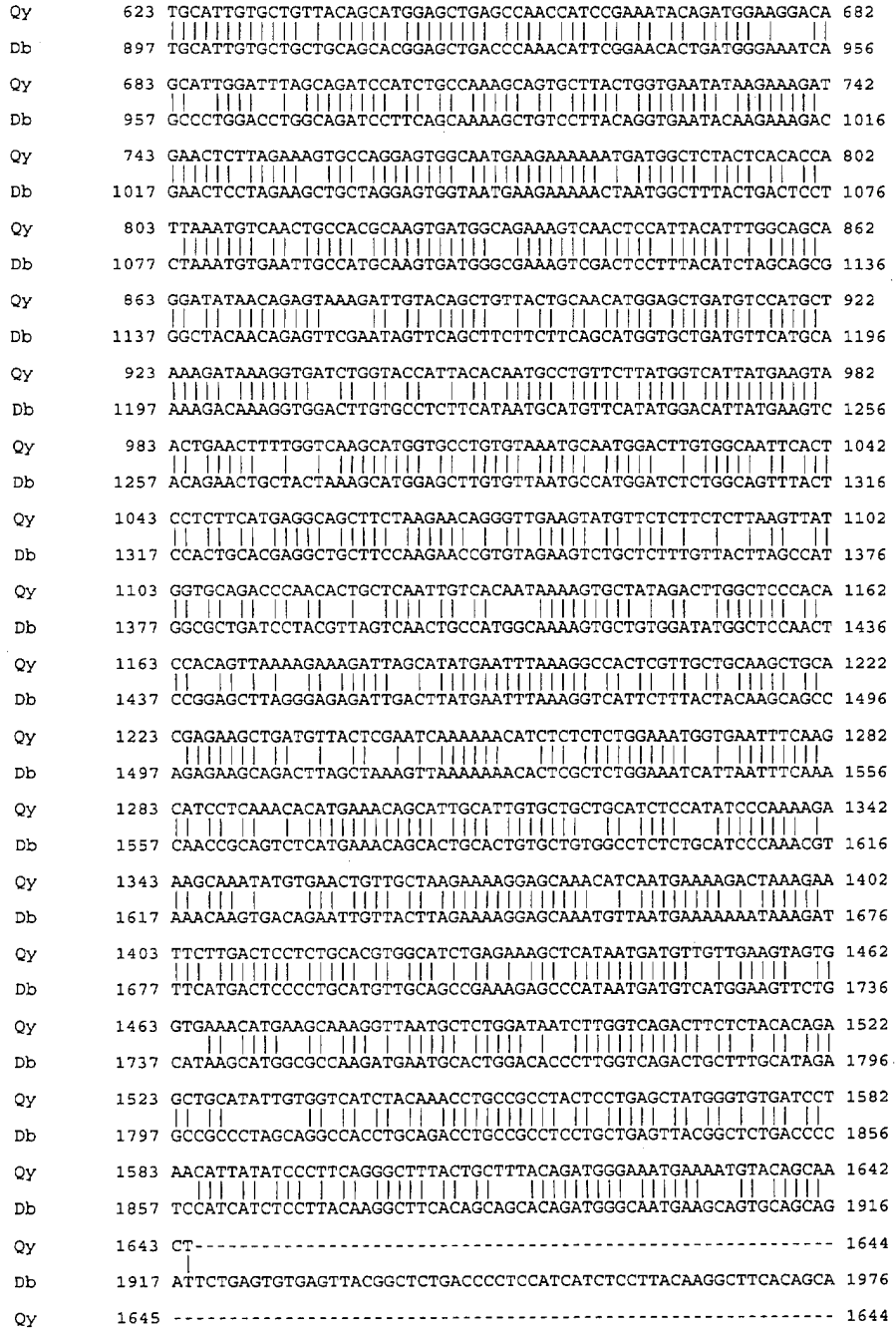
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

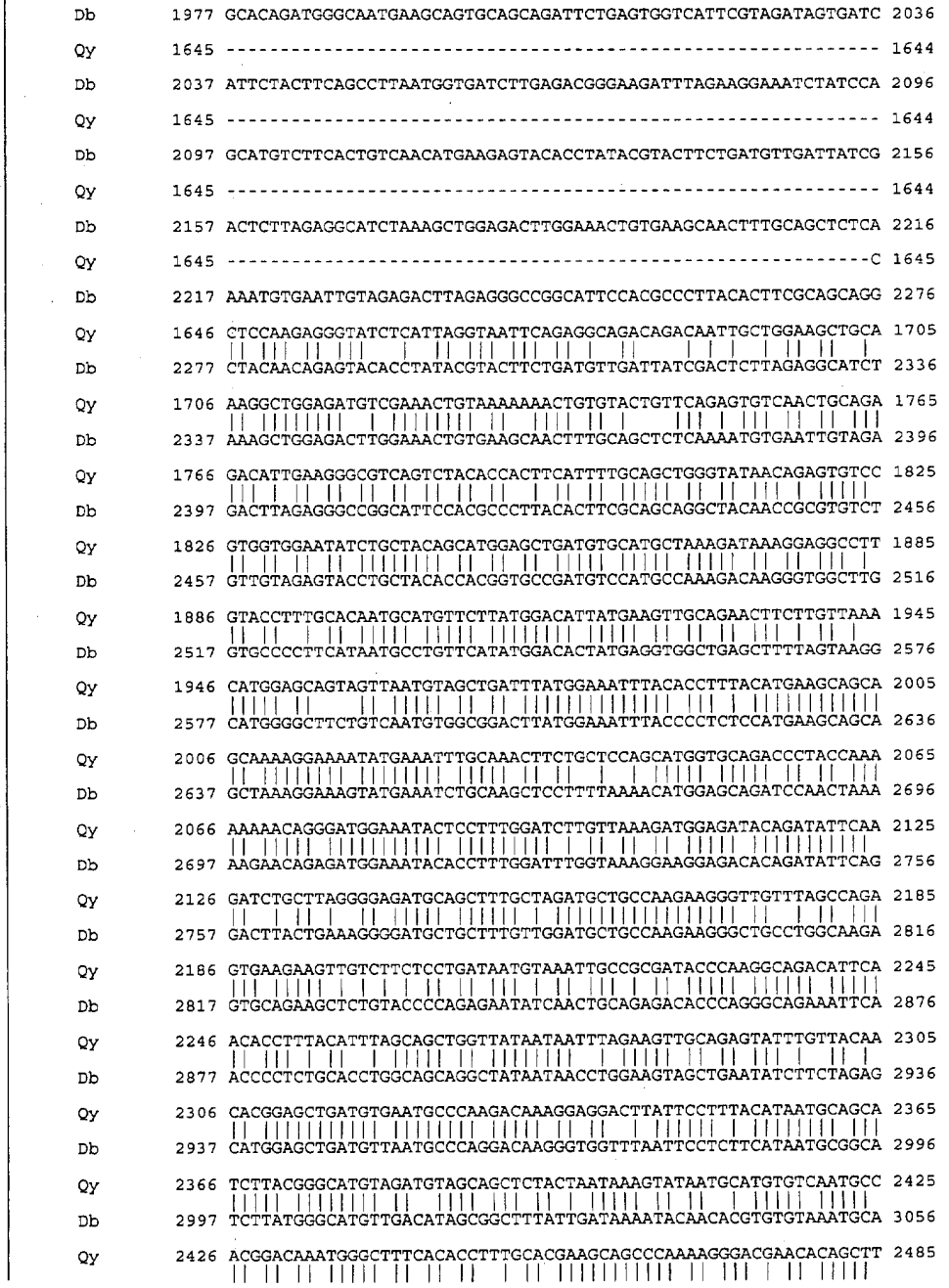
Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

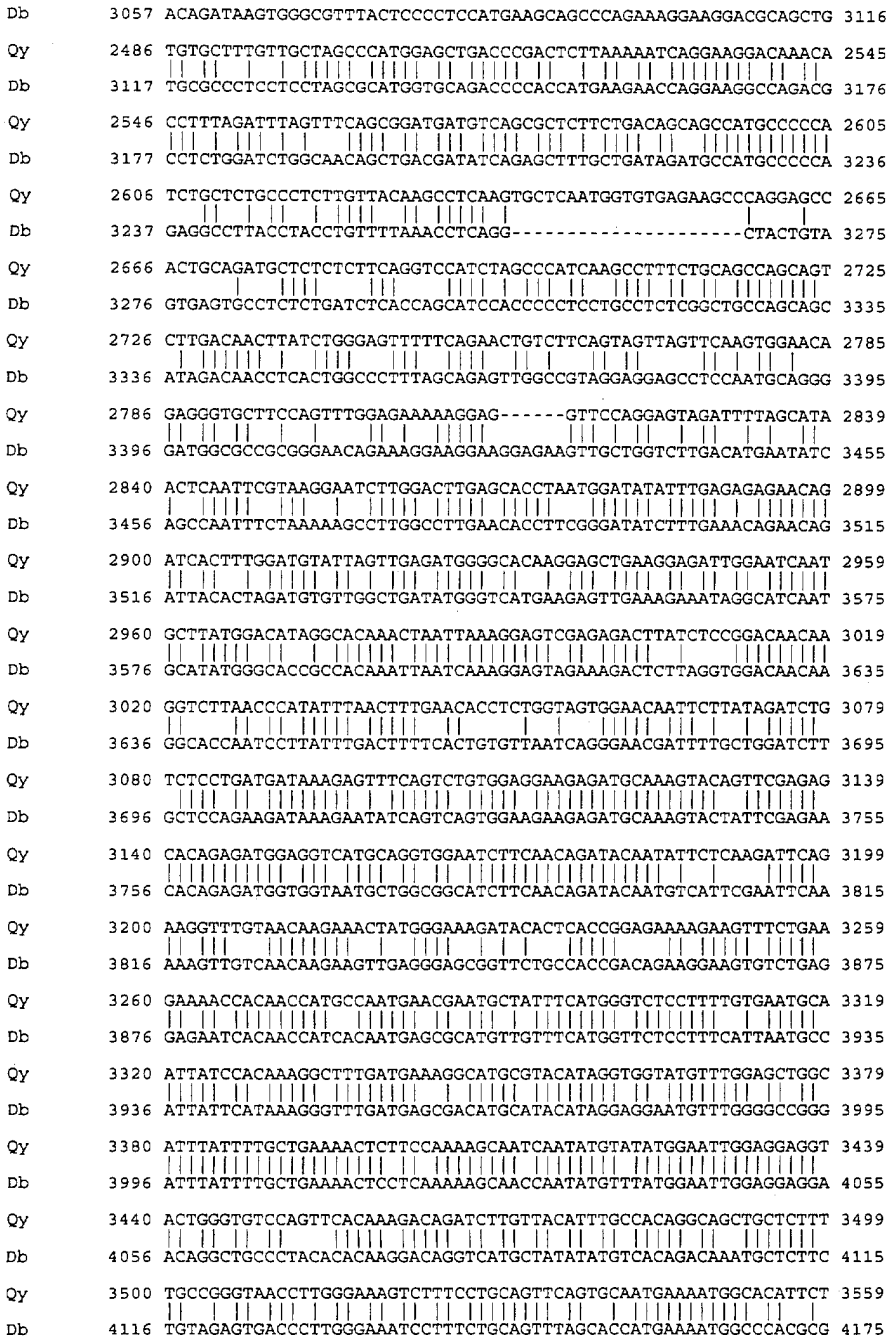
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

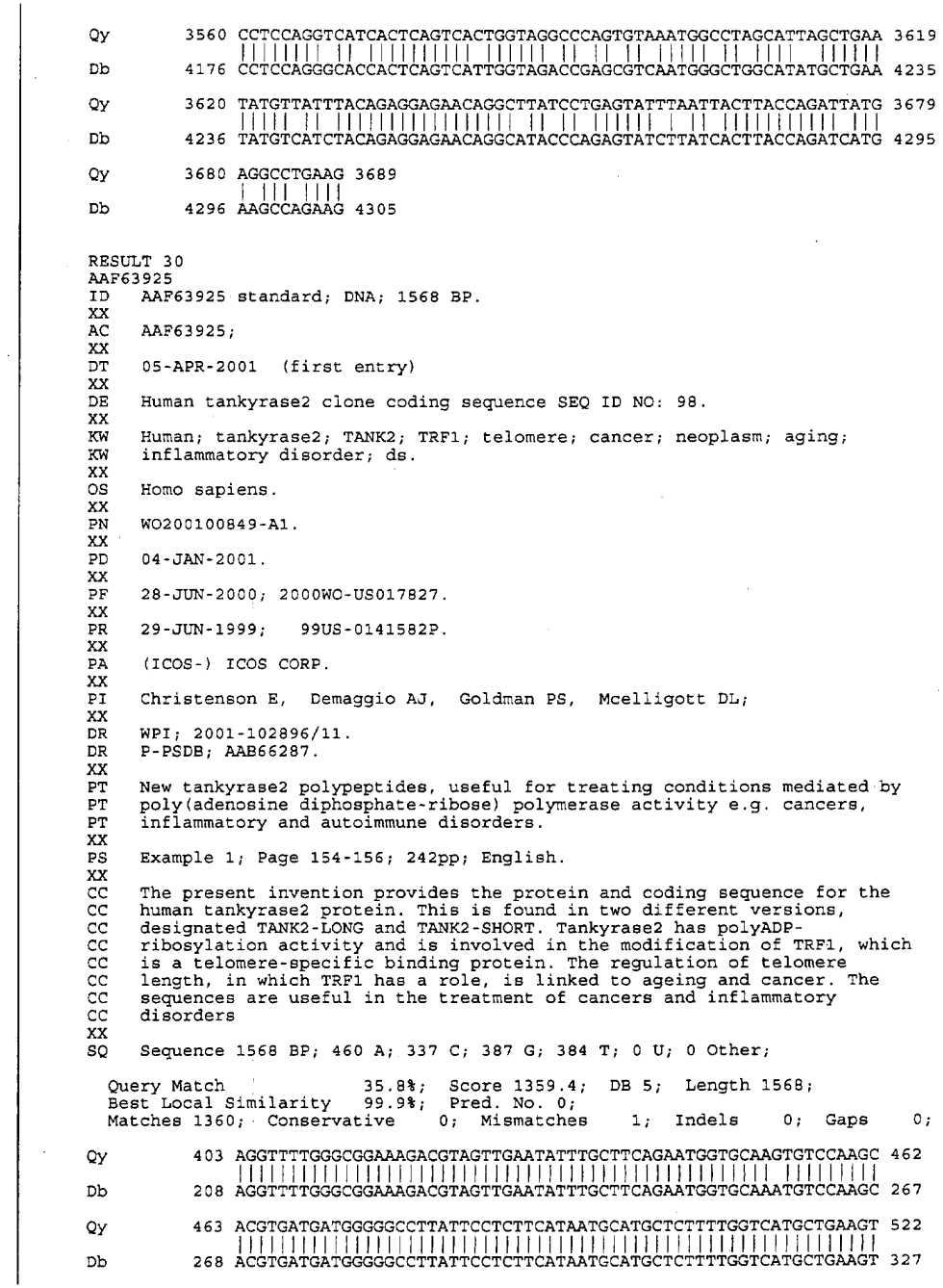
Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896
```









RESULT 30
AAF63925
ID AAF63925 standard; DNA; 1568 BP.
XX
AC AAF63925;
XX
DT 05-APR-2001 (first entry)
XX
DE Human tankyrase2 clone coding sequence SEQ ID NO: 98.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66287.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 1; Page 154-156; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 1568 BP; 460 A; 337 C; 387 G; 384 T; 0 U; 0 Other;

Query Match 35.8%; Score 1359.4; DB 5; Length 1568;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1360; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db       328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy       583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db       388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy       643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db       448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy       703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db       508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy       763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db       568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy       823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db       628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy       883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db       688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy       943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db       748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy      1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db       808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy      1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db       868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

Qy      1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db       928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

Qy      1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db       988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

Qy      1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db      1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

Qy      1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db      1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

Qy      1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db      1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227

Qy      1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db      1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287

Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db      1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347

Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db      1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407

Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db      1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467

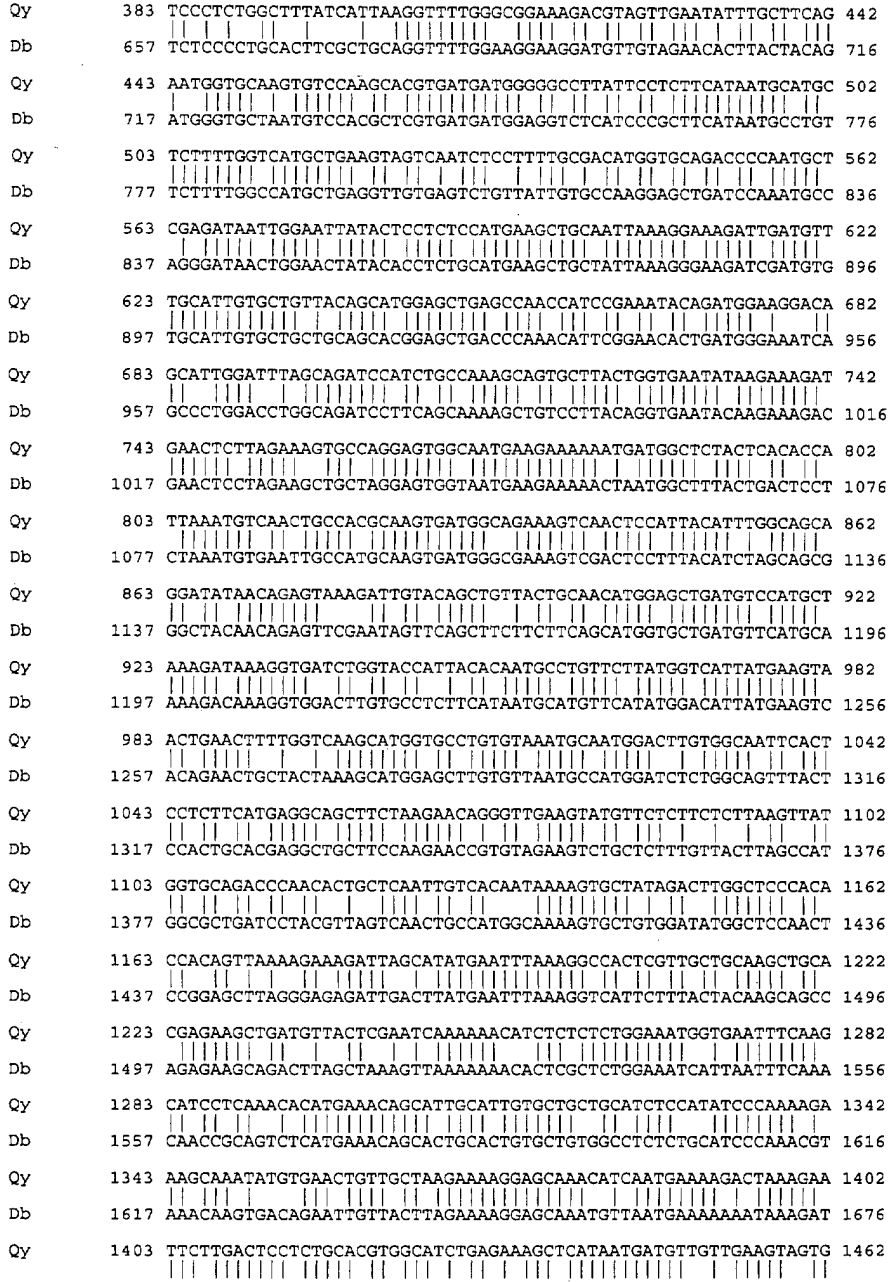
Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db      1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527

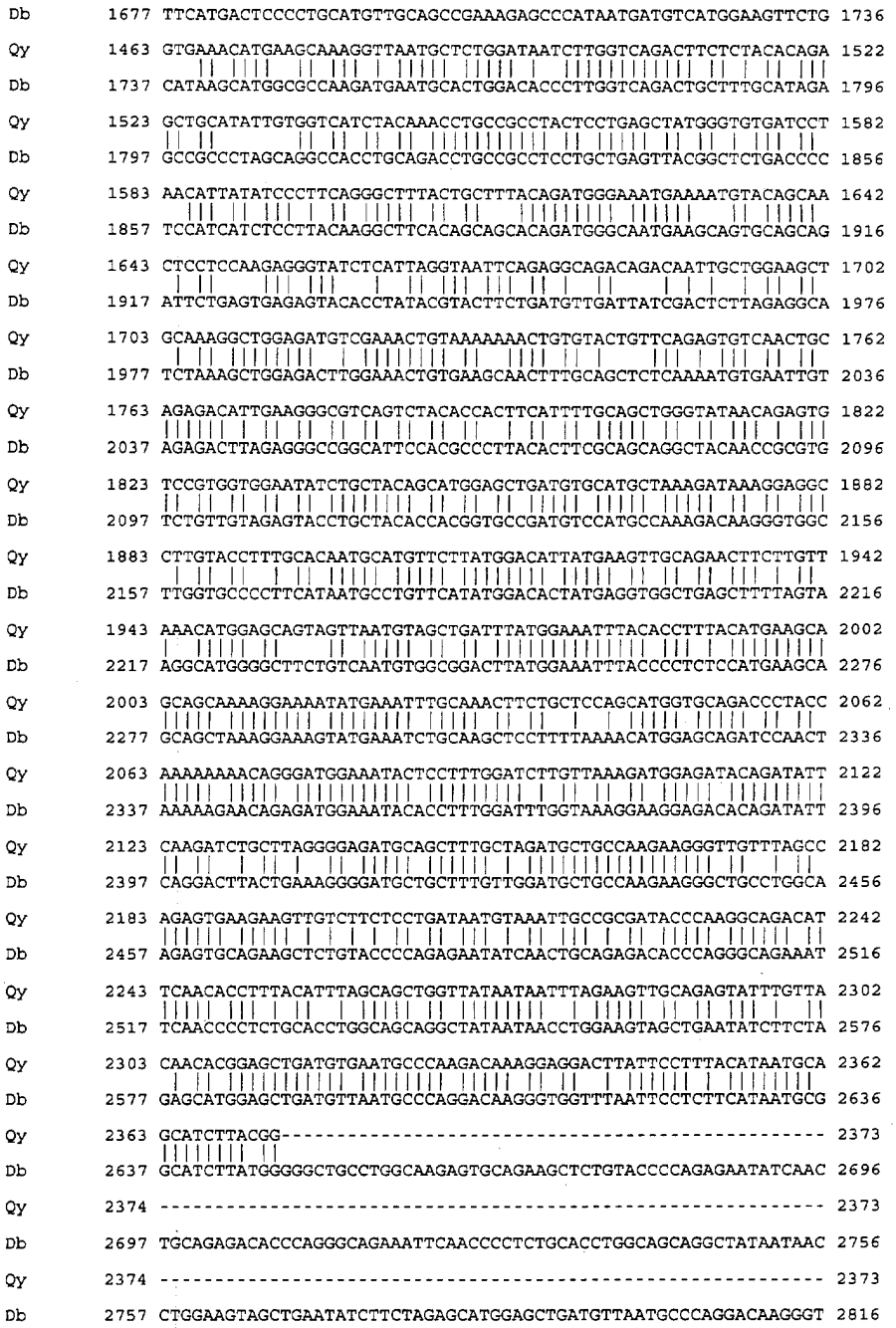
Qy      1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1763
Db      1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1568
```

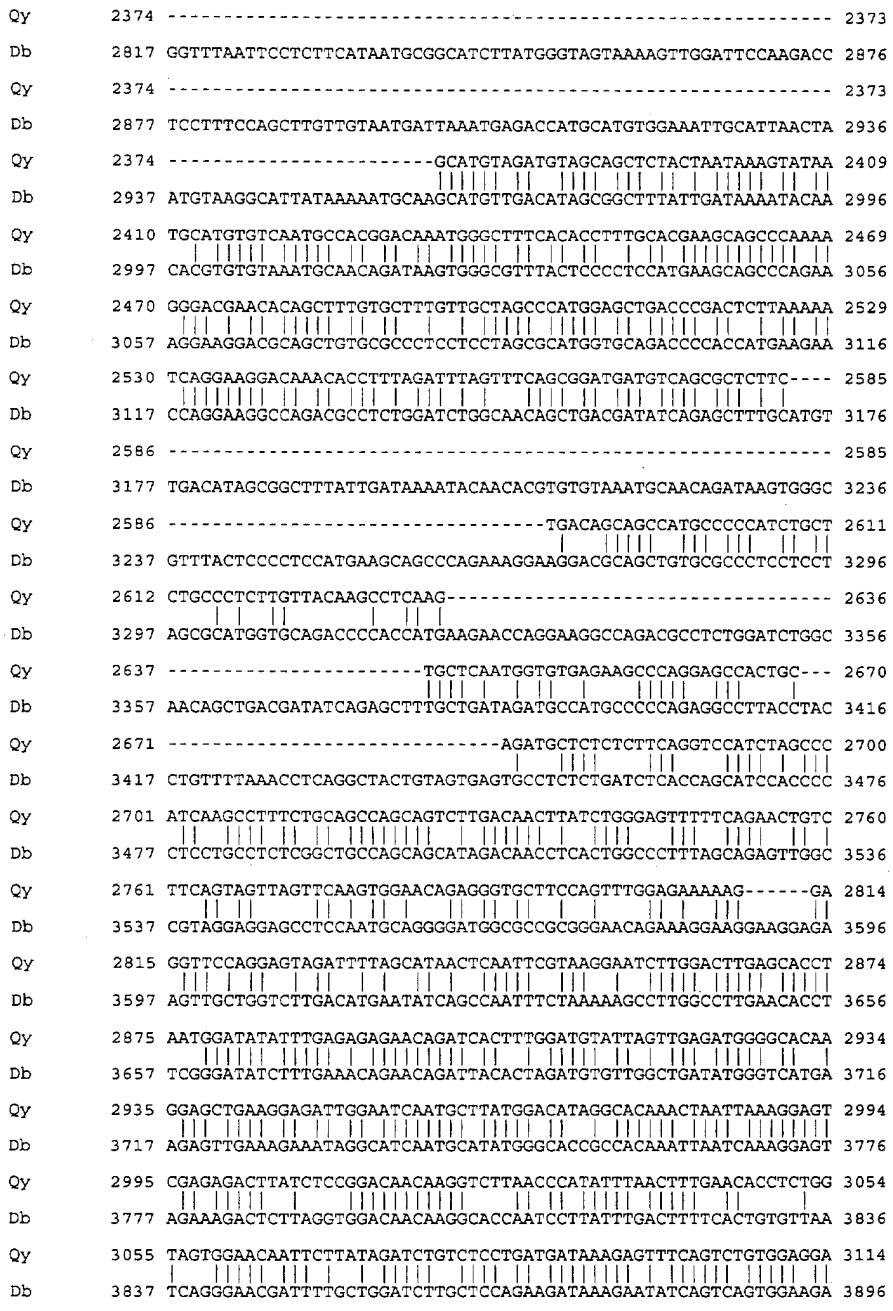
RESULT 31
AAZ29633
ID AAZ29633 standard; DNA; 4657 BP.
XX
AC AAZ29633;
XX
DT 22-MAR-2000 (first entry)
XX
DE Human truncated tankyrase-2 DNA.
XX
KW Human tankyrase; TRF1; telomeric repeat binding factor-1;
KW recombinant tankyrase; telomere length homeostasis; skin atrophy;
KW macular degeneration; atherosclerosis; gene therapy;
KW ataxia telangiectasia; drug screening; ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   6. .2855
FT   /*tag= a
FT   /product= "Human truncated tankyrase"
XX
PN WO9964606-A1.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-US012968.
XX
PR 10-JUN-1998; 98US-00095225.
PR 17-AUG-1998; 98US-00135233.
PR 19-NOV-1998; 98US-00196387.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI De Lange T, Smith S;
XX
DR WPI; 2000-116549/10.
DR P-PSDB; AAY44404.
XX
PT New nucleic acid encoding vertebrate tankyrase, a regulator of telomere
PT length, used to identify modulators, e.g. for inhibiting growth of
PT cancer.
XX
PS Disclosure; Page 123-125; 133pp; English.
XX
CC The present sequence encodes a truncated human tankyrase, a TRF1 binding
CC protein. This DNA comprises the coding sequence for amino acid residues 1
CC -881 of human tankyrase. Tankyrase is involved in regulation of telomere
CC length through TRF1 and may directly modulate the effect of TRF1. This
CC gene is used to produce recombinant tankyrase, or its fragments, and
CC these are used to identify specific modulators which are potential drugs
CC to control telomere length homeostasis for countering telomere shortening
CC associated with aging (e.g. atrophy of skin, macular degeneration or
CC atherosclerosis) or diseases such as ataxia telangiectasia. This can also
CC be used to identify homologous genes and in gene therapy (including
CC antisense or ribozyme expression). Transgenic animals having at least one
CC disrupted endogenous allele for tankyrase are used in drug screening
XX
SQ Sequence 4657 BP; 1268 A; 1162 C; 1160 G; 1067 T; 0 U; 0 Other;

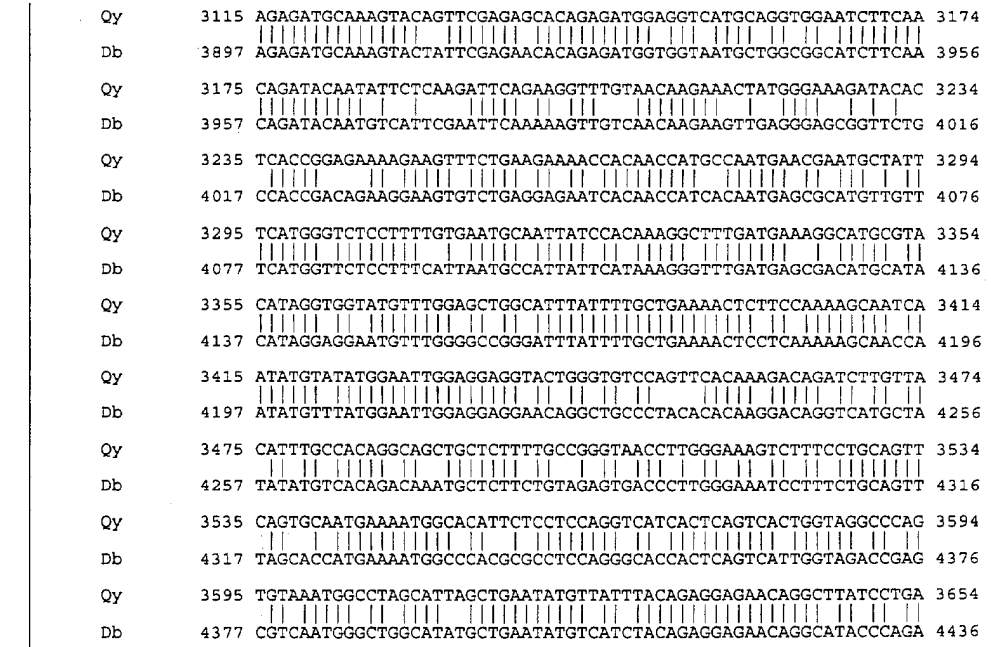
  Query Match 34.2%; Score 1297.8; DB 3; Length 4657;

Best Local Similarity 62.9%; Pred. No. 0;
Matches 2400; Conservative 0; Mismatches 907; Indels 508; Gaps 5;





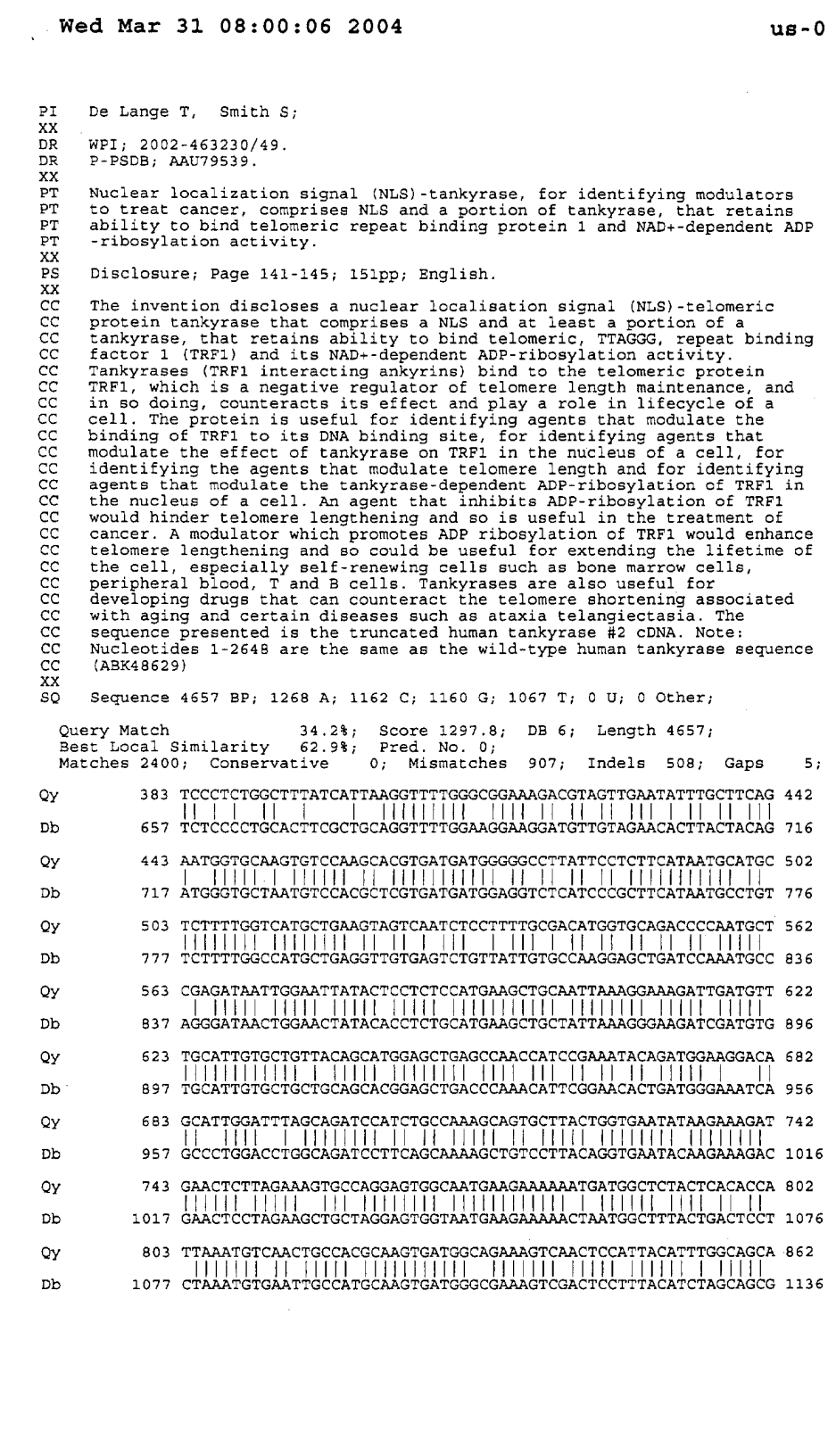


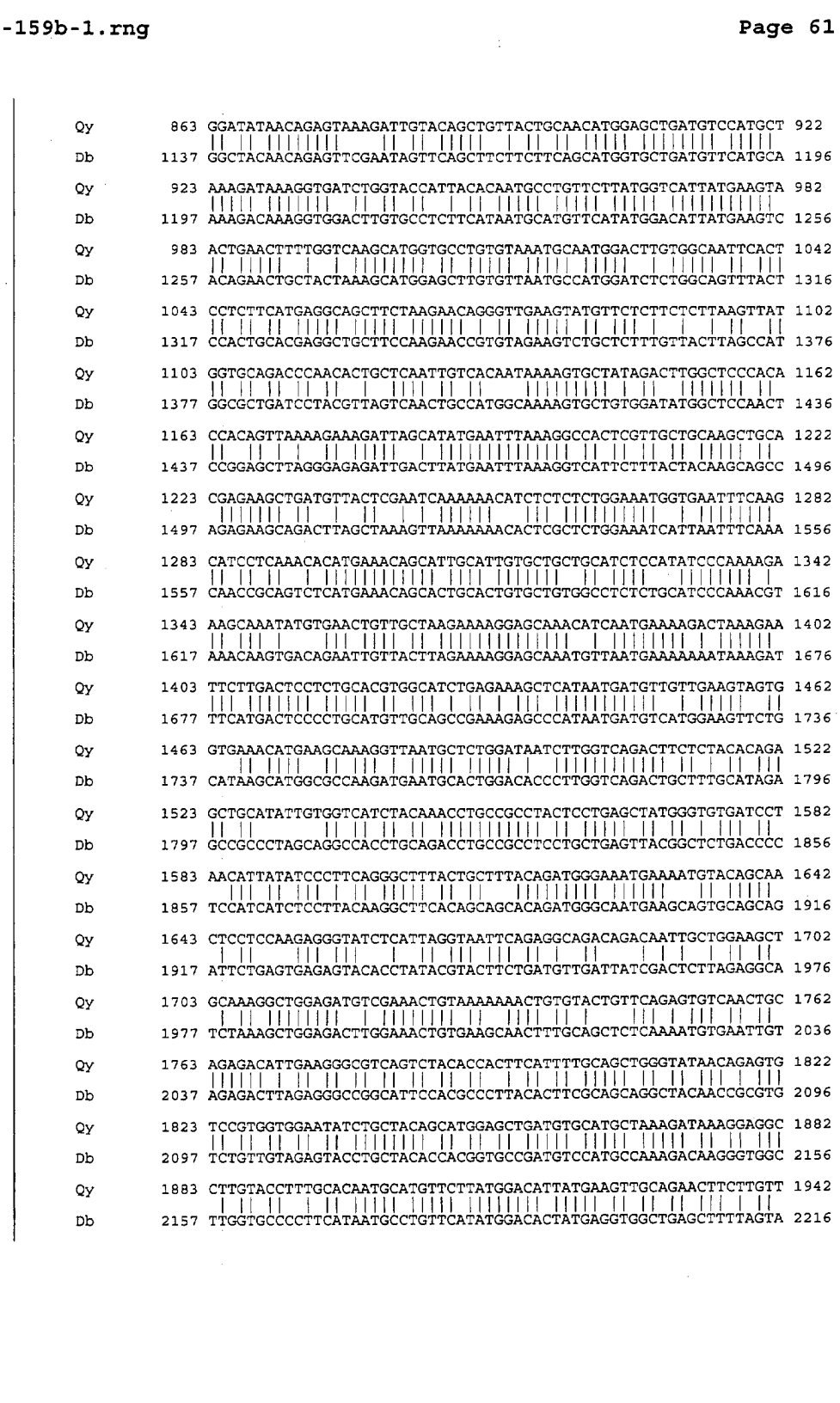


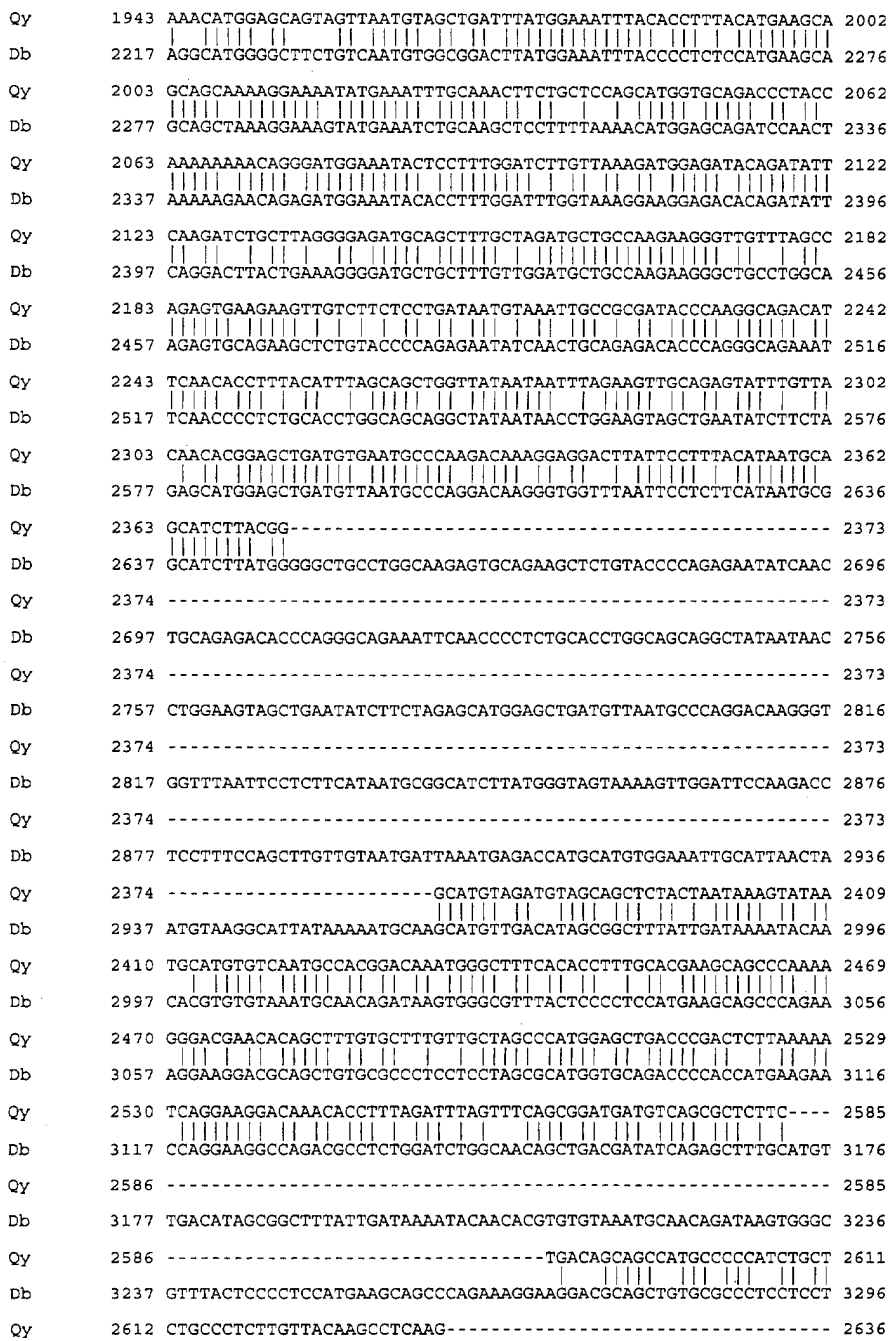
RESULT 32
ABK48635
ID ABK48635 standard; cDNA; 4657 BP.
XX
AC ABK48635;
XX
DT 12-AUG-2002 (first entry)
XX
DE Truncated human tankyrase #2 cDNA.
XX
KW Human; tankyrase; TRF1 interacting ankyrin; ss; gene; cytostatic;
KW ADP ribosylation; telomere; nuclear localisation signal; NLS;
KW telomeric repeat binding factor 1; TRF1; cancer; cell cycle;
KW drug development; aging; ataxia telangiectasia.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 6. .2855
FT /*tag= a
FT /product= "Truncated tankyrase #2"
XX
PN WO200229017-A2.
XX
PD 11-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029931.
XX
PR 03-OCT-2000; 2000US-00678411.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX

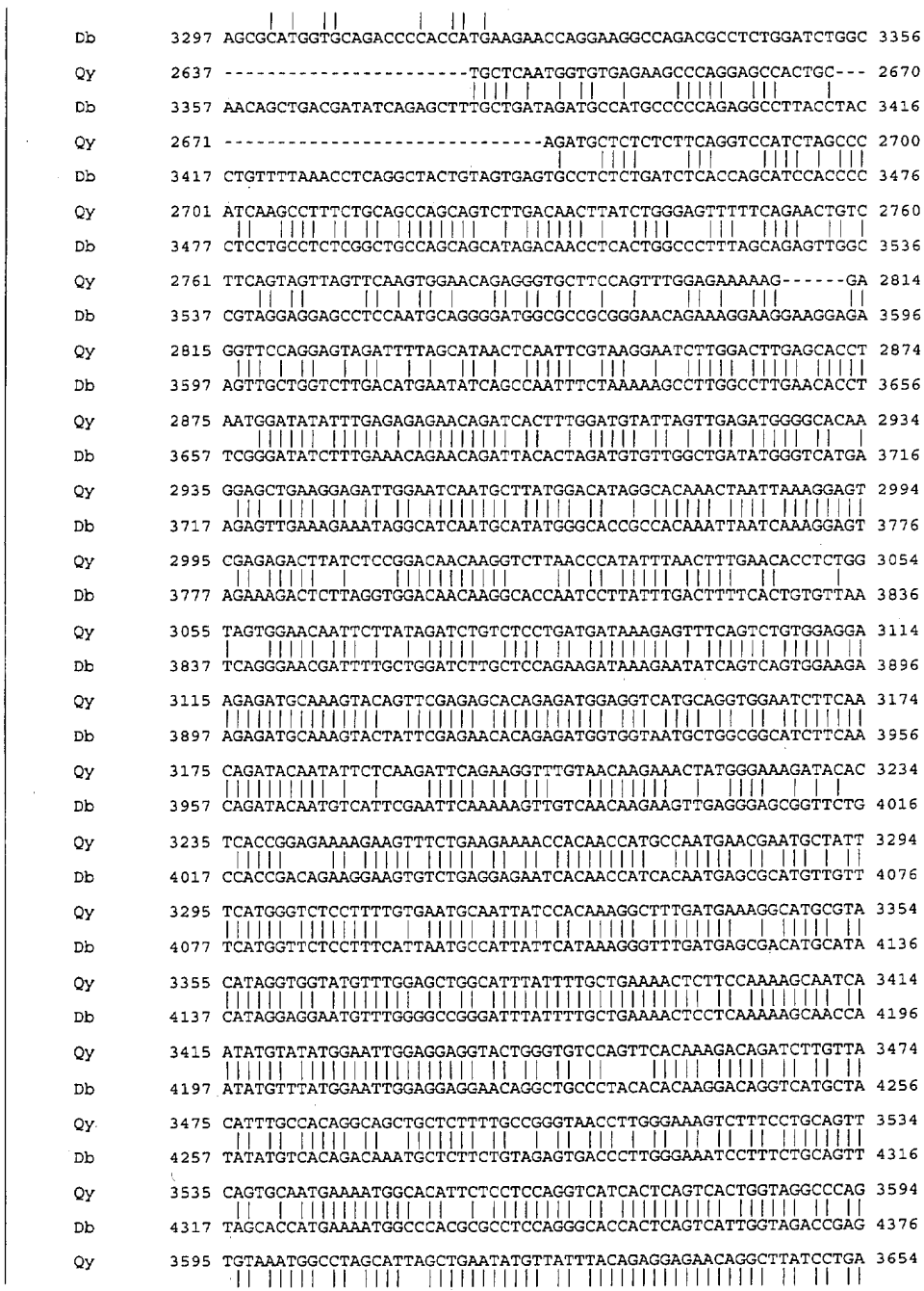
PI De Lange T, Smith S;
XX
DR WPI; 2002-463230/49.
DR P-PSDB; AAU79539.
XX
PT Nuclear localization signal (NLS)-tankyrase, for identifying modulators
PT to treat cancer, comprises NLS and a portion of tankyrase, that retains
PT ability to bind telomeric repeat binding protein 1 and NAD+-dependent ADP
PT -ribosylation activity.
XX
PS Disclosure; Page 141-145; 151pp; English.
XX
CC The invention discloses a nuclear localisation signal (NLS)-telomeric
CC protein tankyrase that comprises a NLS and at least a portion of a
CC tankyrase, that retains ability to bind telomeric, TTAGGG, repeat binding
CC factor 1 (TRF1) and its NAD+-dependent ADP-ribosylation activity.
CC Tankyrases (TRF1 interacting ankyrins) bind to the telomeric protein
CC TRF1, which is a negative regulator of telomere length maintenance, and
CC in so doing, counteracts its effect and play a role in lifecycle of a
CC cell. The protein is useful for identifying agents that modulate the
CC binding of TRF1 to its DNA binding site, for identifying agents that
CC modulate the effect of tankyrase on TRF1 in the nucleus of a cell, for
CC identifying the agents that modulate telomere length and for identifying
CC agents that modulate the tankyrase-dependent ADP-ribosylation of TRF1 in
CC the nucleus of a cell. An agent that inhibits ADP-ribosylation of TRF1
CC would hinder telomere lengthening and so is useful in the treatment of
CC cancer. A modulator which promotes ADP ribosylation of TRF1 would enhance
CC telomere lengthening and so could be useful for extending the lifetime of
CC the cell, especially self-renewing cells such as bone marrow cells,
CC peripheral blood, T and B cells. Tankyrases are also useful for
CC developing drugs that can counteract the telomere shortening associated
CC with aging and certain diseases such as ataxia telangiectasia. The
CC sequence presented is the truncated human tankyrase #2 cDNA. Note:
CC Nucleotides 1-2648 are the same as the wild-type human tankyrase sequence
CC (ABK48629)
XX
SQ Sequence 4657 BP; 1268 A; 1162 C; 1160 G; 1067 T; 0 U; 0 Other;

Query Match 34.2%; Score 1297.8; DB 6; Length 4657;
Best Local Similarity 62.9%; Pred. No. 0;
Matches 2400; Conservative 0; Mismatches 907; Indels 508; Gaps 5;









```
Db       4377 CGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATACCCAGA 4436

Qy       3655 GTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3689
              |||| | || |||||||||||| |||| ||| ||||
Db       4437 GTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4471


RESULT 33
ABA08564
ID   ABA08564 standard; cDNA; 1466 BP.
XX
AC   ABA08564;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human Grb7 effector protein homologue-encoding cDNA, SEQ ID NO:340.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US003800.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11320.
XX
PT   Human proteins and DNA encoding sequences useful for preventing, treating
PT   or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT   and cancer.
XX
PS   Claim 1; Page 453; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention
XX
SQ   Sequence 1466 BP; 468 A; 287 C; 321 G; 390 T; 0 U; 0 Other;

  Query Match             32.7%;  Score 1240.4;  DB 4;  Length 1466;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1247;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1023 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 1082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 60

Qy       1083 GTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 1142
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         61 GTTCTCTTCTCTTAAGTTATGGGGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 120

Qy       1143 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 1202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 180

Qy       1203 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCTC 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAAACATCTCTCTC 240

Qy       1263 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 1322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 300

Qy       1323 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 1382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 360

Qy       1383 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 1442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 420

Qy       1443 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 1502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 480

Qy       1503 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 1562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 540

Qy       1563 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 1622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 600

Qy       1623 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 1682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 660

Qy       1683 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTA 1742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTA 720
```

```
Qy       1743 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 780

Qy       1803 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 840

Qy       1863 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 1922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 900

Qy       1923 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 1982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 960

Qy       1983 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 1020

Qy       2043 AGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 2102
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 1080

Qy       2103 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 1140

Qy       2163 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 2222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 1200

Qy       2223 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
             ||||||||||||||||||||||||||||||||||||||||| |   |  |  || ||
Db       1201 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGGTAAGTGAATGAAGTT 1258
```

RESULT 34
AAC89701
ID AAC89701 standard; cDNA; 1242 BP.
XX
AC AAC89701;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human adult T-cell leukaemia cDNA clone HOM-TALL1-5(5').
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR P-PSDB; AAB48571.
XX
PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 1; Page 74-75; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 1242 BP; 391 A; 249 C; 269 G; 328 T; 0 U; 5 Other;

Query Match 30.9%; Score 1172.2; DB 4; Length 1242;
Best Local Similarity 97.8%; Pred. No. 2.1e-314;
Matches 1218; Conservative 0; Mismatches 23; Indels 5; Gaps 3;

```
Qy        510 GTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATA 569
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATA 60

Qy        570 ATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTG 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTG 120

Qy        630 TGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGG 689
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGG 180

Qy        690 ATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCT 749
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCT 240

Qy        750 TAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATG 809
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAGAAAGTGCCAGGNGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATG 300

Qy        810 TCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATA 869
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATA 360

Qy        870 ACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATA 929
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATA 420

Qy        930 AAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAAC 989
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAAC 480

Qy        990 TTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTC 1049
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTC 540

Qy       1050 ATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAG 1109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAG 600

Qy       1110 ACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGT 1169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGT 660

Qy       1170 TAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAG 1229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAG 720

Qy       1230 CTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTC 1289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        721 CTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTC 780

Qy       1290 AAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAA 1349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAA 840

Qy       1350 TATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGA 1409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGA 900

Qy       1410 CTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAAC 1469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAAC 960

Qy       1470 ATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAT 1529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAT 1020

Qy       1530 ATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTA 1589
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATTGTGGTCATCTACACACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTA 1080

Qy       1590 TATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCC 1649
              |||||||||||||||||||||||||||||||||||    |||||||||||| || |
Db       1081 TATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAA--TGAAATGTACAGCAACTNCTNC 1138

Qy       1650 AAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAA-TTGCTGGAAGCTGCAAAG 1708
              |||||||||||||||||||||||||||||||||||||||   || |||| |       ||
Db       1139 AAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACANTTTNCTGGTAAGCTGCAAA 1198

Qy       1709 GCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTG 1754
              |   ||||||||||      |||||||||||||||||||||||||
Db       1199 GGCTGAGATGTCGAA--CTCTAAAAAACTGTGTACTGTTCAGAGTG 1242


RESULT 35
AAF63990
ID   AAF63990 standard; DNA; 3308 BP.
XX
AC   AAF63990;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human tankyrase2 expression plasmid coding sequence SEQ ID NO: 177.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;  99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66301.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 7; Page 232-236; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 3308 BP; 958 A; 730 C; 900 G; 720 T; 0 U; 0 Other;

  Query Match             30.8%;  Score 1168.4;  DB 5;  Length 3308;
  Best Local Similarity   99.9%;  Pred. No. 3.5e-313;
  Matches 1169;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2534 GAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCA 2593
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       2101 GAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCA 2160

Qy       2594 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2220

Qy       2654 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2280

Qy       2714 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2340

Qy       2774 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2400

Qy       2834 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2460

Qy       2894 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2520

Qy       2954 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 3013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 2580

Qy       3014 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 3073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 2640

Qy       3074 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 3133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 2700

Qy       3134 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 3193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 2760

Qy       3194 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 3253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 2820

Qy       3254 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 3313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 2880

Qy       3314 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 3373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 2940

Qy       3374 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3000
```

```
Qy       3434 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3060

Qy       3494 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3120

Qy       3554 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3180

Qy       3614 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3240

Qy       3674 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
             |||||||||||||||||||||||||||||
Db       3241 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3270
```

RESULT 36
ACA57422
ID ACA57422 standard; cDNA; 1121 BP.
XX
AC ACA57422;
XX
DT 10-JUN-2003 (first entry)
XX
DE Human adipocyte Selected Interacting domain, SID, cDNA #509.
XX
KW Human; ss; gene; prey; adipocyte; SID; selected interacting domain;
KW anorectic; antidiabetic; protein-protein interaction; diabetes;
KW yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS Homo sapiens.
XX
PN WO200286122-A2.
XX
PD 31-OCT-2002.
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Daviet L;
XX
DR WPI; 2003-103412/09.
DR P-PSDB; ABU70878.
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
PS Claim 7; Page 276; 382pp; English.
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM)) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence encodes a SID
CC (prey) protein of the invention
XX
SQ Sequence 1121 BP; 354 A; 228 C; 241 G; 298 T; 0 U; 0 Other;

Query Match 29.5%; Score 1119.4; DB 7; Length 1121;
Best Local Similarity 99.9%; Pred. No. 8.1e-300;
Matches 1120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        439 TCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGC 498
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCAGAATGGTGCAAATGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGC 60

Qy        499 ATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAA 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAA 120

Qy        559 TGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGA 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGA 180

Qy        619 TGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAG 240

Qy        679 GACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAA 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAA 300

Qy        739 AGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCAC 798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCAC 360

Qy        799 ACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGC 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGC 420

Qy        859 AGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCA 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCA 480

Qy        919 TGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGA 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGA 540

Qy        979 AGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATT 1038
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATT 600

Qy       1039 CACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAG 1098
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAG 660

Qy       1099 TTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCC 1158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCC 720

Qy       1159 CACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGC 1218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGC 780

Qy       1219 TGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTT 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTT 840

Qy       1279 CAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAA 1338
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAA 900

Qy       1339 AAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAA 1398
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAA 960

Qy       1399 AGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGT 1458
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGT 1020

Qy       1459 AGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACA 1518
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACA 1080

Qy       1519 CAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTAC 1559
                |||||||||||||||||||||||||||||||||||||||
Db       1081 CAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTAC 1121


RESULT 37
AAF63973
ID   AAF63973 standard; DNA; 1103 BP.
XX
AC   AAF63973;
XX
DT   05-APR-2001  (first entry)
XX
DE   Human TRF1 TANK2 binding domain coding sequence SEQ ID NO: 159.
XX
KW   Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW   inflammatory disorder; ds.
XX
OS   Homo sapiens.
XX
PN   WO200100849-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017827.
XX
PR   29-JUN-1999;   99US-0141582P.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Christenson E,  Demaggio AJ,  Goldman PS,  Mcelligott DL;
XX
DR   WPI; 2001-102896/11.
DR   P-PSDB; AAB66300.
XX
PT   New tankyrase2 polypeptides, useful for treating conditions mediated by
PT   poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT   inflammatory and autoimmune disorders.
XX
PS   Example 6; Page 226-228; 242pp; English.
XX
CC   The present invention provides the protein and coding sequence for the
CC   human tankyrase2 protein. This is found in two different versions,
CC   designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC   ribosylation activity and is involved in the modification of TRF1, which
CC   is a telomere-specific binding protein. The regulation of telomere
CC   length, in which TRF1 has a role, is linked to ageing and cancer. The
CC   sequences are useful in the treatment of cancers and inflammatory
CC   disorders
XX
SQ   Sequence 1103 BP; 352 A; 219 C; 239 G; 293 T; 0 U; 0 Other;

  Query Match           28.6%;  Score 1084.8;  DB 5;  Length 1103;
  Best Local Similarity   99.8%;  Pred. No. 3.3e-290;
  Matches 1086;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1037 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 1096
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 68

Qy       1097 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 1156
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 128

Qy       1157 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 1216
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 188

Qy       1217 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 1276
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 248

Qy       1277 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 1336
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 308

Qy       1337 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 1396
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 368

Qy       1397 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 1456
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 428

Qy       1457 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 1516
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 488

Qy       1517 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 1576
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 548

Qy       1577 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 1636
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 608

Qy       1637 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 1696
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 668

Qy       1697 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTC 1756
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTC 728

Qy       1757 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 1816
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 788

Qy       1817 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 1876
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 848

Qy       1877 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 1936
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 908

Qy       1937 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 1996
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 968

Qy       1997 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 2056
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 1028

Qy       2057 CCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 2116
                ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 CCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 1088

Qy       2117 GATATTCA 2124
                |||||| |
Db       1089 GATATTTA 1096
```

```
RESULT 38
AAH16624
ID   AAH16624 standard; cDNA; 3006 BP.
XX
AC   AAH16624;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:15732.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-00116126.
XX
PR   29-JUL-1999;   99JP-00248036.
PR   27-AUG-1999;   99JP-00300253.
PR   11-JAN-2000; 2000JP-00118776.
PR   02-MAY-2000; 2000JP-00183767.
PR   09-JUN-2000; 2000JP-00241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT   length cDNAs defined in the specification, and for the detection and/or
PT   diagnosis of the abnormality of the proteins encoded by the full-length
PT   cDNAs.
XX
PS   Claim 8; SEQ ID NO 15732; 2537pp + Sequence Listing; English.
XX
CC   The present invention describes primer sets for synthesising 5602 full-
CC   length cDNAs defined in the specification. Where a primer set comprises:
CC   (a) an oligo-dT primer and an oligonucleotide complementary to the
CC   complementary strand of a polynucleotide which comprises one of the 5602
CC   nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in the
CC   specification. The primer sets can be used in antisense therapy and in
CC   gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC   represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC   oligonucleotides, all of which are used in the exemplification of the
CC   present invention
XX
SQ   Sequence 3006 BP; 892 A; 505 C; 584 G; 1025 T; 0 U; 0 Other;

  Query Match           28.4%;  Score 1076.8;  DB 4;  Length 3006;
  Best Local Similarity  99.8%;  Pred. No. 9.4e-288;
  Matches 1078;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2715 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 60

Qy       2775 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 120

Qy       2835 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 180

Qy       2895 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 240

Qy       2955 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 300

Qy       3015 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 3074
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 360

Qy       3075 ATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 3134
             ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         361 ATCTGTCTCCTGATGATAAAGAGTTTCCGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 420

Qy       3135 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 3194
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 480

Qy       3195 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 540

Qy       3255 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 600

Qy       3315 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 660

Qy       3375 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3434
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 720

Qy       3435 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3494
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 780

Qy       3495 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 3554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 840

Qy       3555 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3614
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 900

Qy       3615 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3674
             ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         901 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTAAATTACTTACCAGA 960

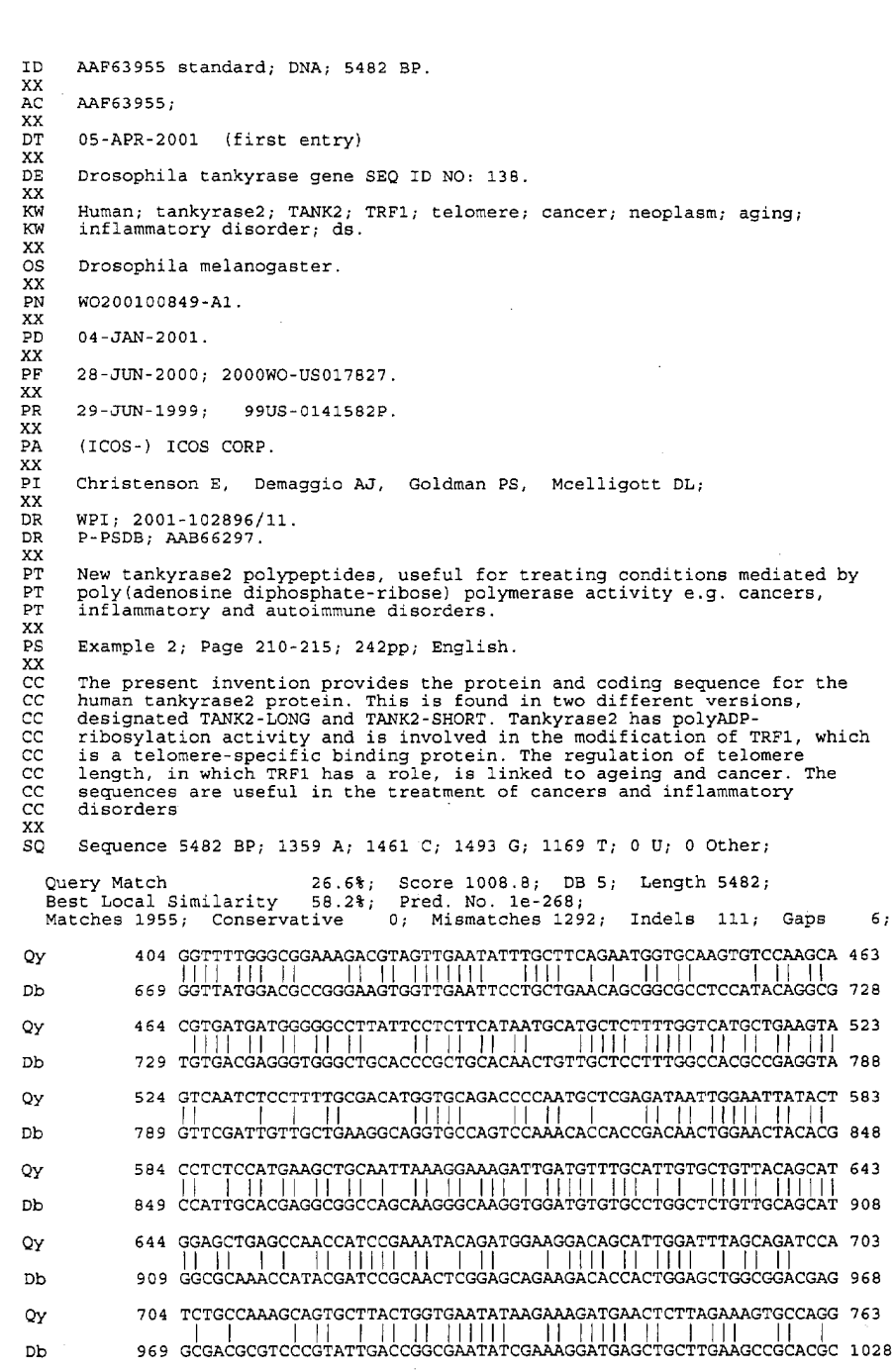
Qy       3675 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3734
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 1020

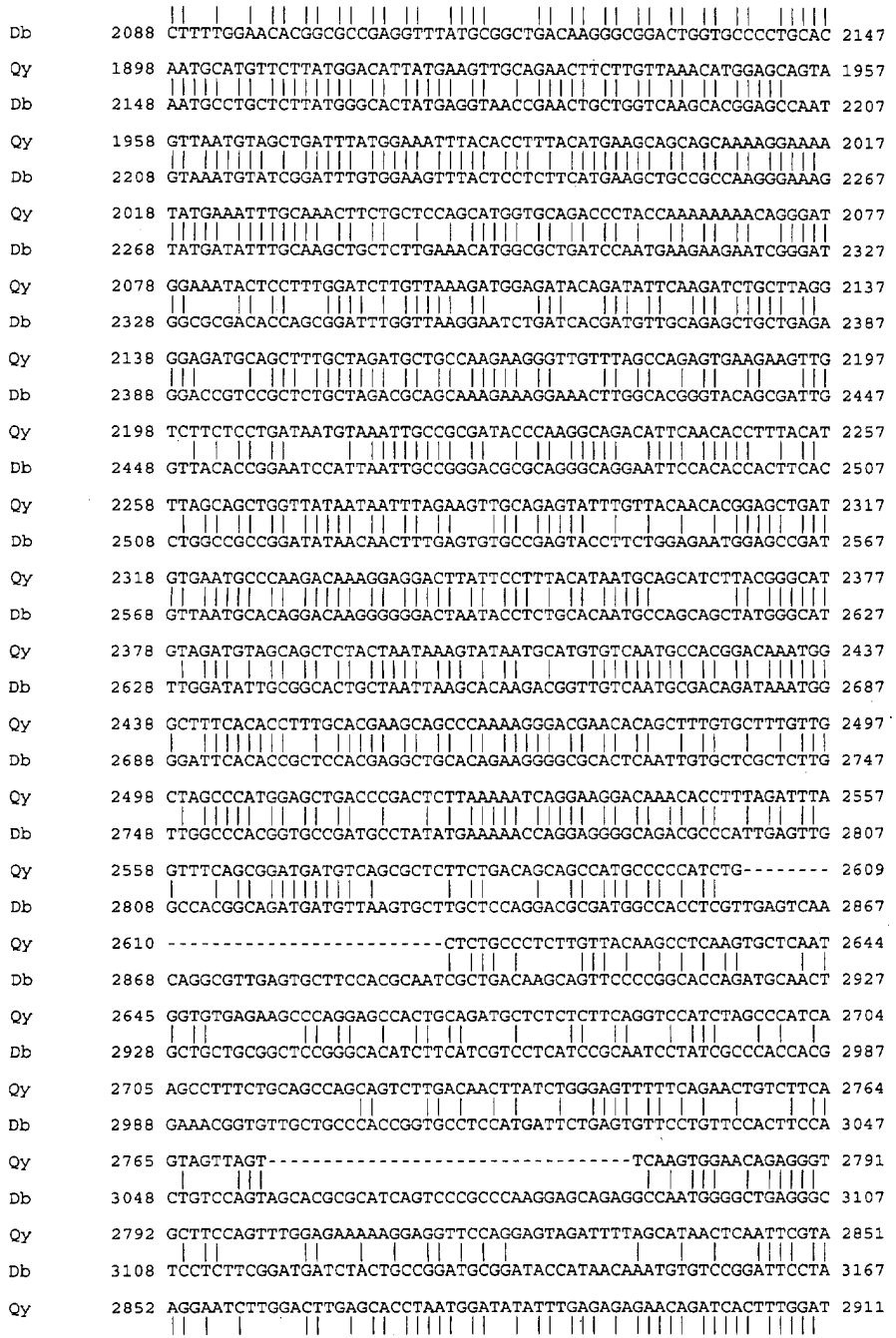
Qy       3735 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 1080

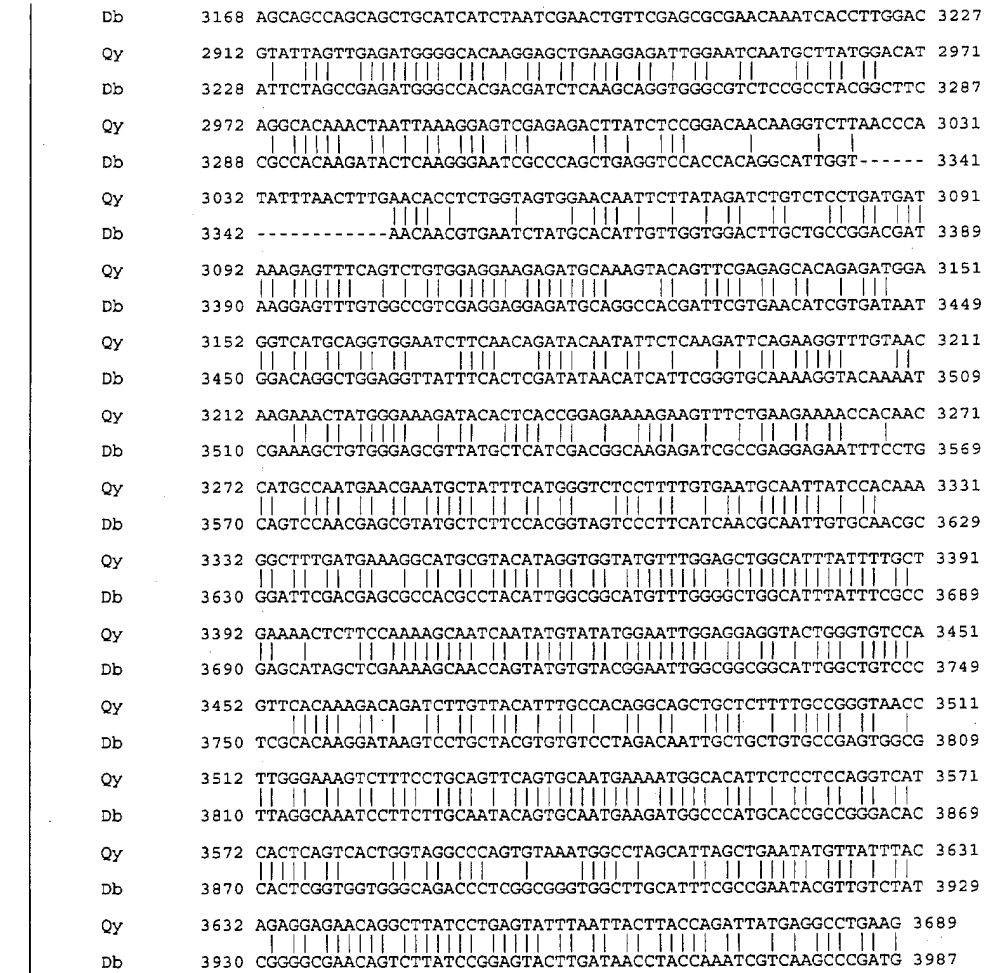
RESULT 39
AAF63955
```

ID AAF63955 standard; DNA; 5482 BP.
XX
AC AAF63955;
XX
DT 05-APR-2001 (first entry)
XX
DE Drosophila tankyrase gene SEQ ID NO: 138.
XX
KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
KW inflammatory disorder; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200100849-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US017827.
XX
PR 29-JUN-1999; 99US-0141582P.
XX
PA (ICOS-) ICOS CORP.
XX
PI Christenson E, Demaggio AJ, Goldman PS, Mcelligott DL;
XX
DR WPI; 2001-102896/11.
DR P-PSDB; AAB66297.
XX
PT New tankyrase2 polypeptides, useful for treating conditions mediated by
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
PT inflammatory and autoimmune disorders.
XX
PS Example 2; Page 210-215; 242pp; English.
XX
CC The present invention provides the protein and coding sequence for the
CC human tankyrase2 protein. This is found in two different versions,
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-
CC ribosylation activity and is involved in the modification of TRF1, which
CC is a telomere-specific binding protein. The regulation of telomere
CC length, in which TRF1 has a role, is linked to ageing and cancer. The
CC sequences are useful in the treatment of cancers and inflammatory
CC disorders
XX
SQ Sequence 5482 BP; 1359 A; 1461 C; 1493 G; 1169 T; 0 U; 0 Other;

Query Match 26.6%; Score 1008.8; DB 5; Length 5482;
Best Local Similarity 58.2%; Pred. No. 1e-268;
Matches 1955; Conservative 0; Mismatches 1292; Indels 111; Gaps 6;







RESULT 40
ABL08155
ID ABL08155 standard; cDNA; 5484 BP.
XX
AC ABL08155;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18947.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.

XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB64052.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 18947; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5484 BP; 1346 A; 1467 C; 1495 G; 1176 T; 0 U; 0 Other;

Query Match 26.6%; Score 1008.8; DB 4; Length 5484;
Best Local Similarity 58.2%; Pred. No. 1e-268;
Matches 1955; Conservative 0; Mismatches 1292; Indels 111; Gaps 6;

```
Qy        404 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCA 463
Db        688 GGTTATGGACGCCGGGAAGTGGTTGAATTCCTGCTGAACAGCGGCGCCTCCATACAGGCG 747

Qy        464 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 523
Db        748 TGTGACGAGGGTGGGCTGCACCCGCTGCACAACTGTTGCTCCTTTGGCCACGCCGAGGTA 807

Qy        524 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 583
Db        808 GTTCGATTGTTGCTGAAGGCAGGTGCCAGTCCAAACACCACCGACAACTGGAACTACACG 867

Qy        584 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 643
Db        868 CCATTGCACGAGGCGGCCAGCAAGGGCAAGGTGGATGTGTGCCTGGCTCTGTTGCAGCAT 927

Qy        644 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 703
Db        928 GGCGCAAACCATACGATCCGCAACTCGGAGCAGAAGACACCACTGGAGCTGGCGGACGAG 987

Qy        704 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 763
Db        988 GCGACGCGTCCCGTATTGACCGGCGAATATCGAAAGGATGAGCTGCTTGAAGCCGCACGC 1047

Qy        764 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 823
Db       1048 TCGGGGGCCGAGGATCGCCTGCTGGCCCTACTCACGCCACTCAATGTCAACTGTCATGCC 1107

Qy        824 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883
Db       1108 AGCGATGGACGACGCTCAACGCCGCTCCATCTGGCAGCGGGCTACAATCGGATCGGCATC 1167

Qy        884 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 943
Db       1168 GTGGAAATTCTGCTGGCCAACGGAGCGGATGTACATGCTAAGGACAAGGGCGGTCTGGTG 1227

Qy        944 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1003

Db       1228 CCGCTGCACAATGCCTGCTCCTACGGACACTTCGATGTGACCAAGCTGCTTATCCAGGCG 1287

Qy       1004 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1063
Db       1288 GGCGCCAATGTCAACGCCAACGATCTGTGGGCCTTTACGCCGCTCCACGAGGCCGCCTCC 1347

Qy       1064 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1123
Db       1348 AAAAGTCGCGTCGAGGTCTGCAGCCTGCTGCTCAGTCGTGGAGCGGATCCCACCCTCCTA 1407

Qy       1124 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1183
Db       1408 AACTGCCACAGCAAGTCGGCCATCGATGCGGCGCCCACCAGGGAGCTGAGAGAGCGGATT 1467

Qy       1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1243
Db       1468 GCCTTTGAATACAAGGGTCACTGCCTGCTGGACGCCTGTCGAAAGTGTGATGTGTCCCGT 1527

Qy       1244 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1303
Db       1528 GCCAAGAAGCTGGTATGCGCAGAGATTGTTAACTTCGTGCATCCATATACAGGAGACACT 1587

Qy       1304 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1363
Db       1588 CCGCTCCACCTGGCCGTTGTCAGTCCGGATGGGAAGCGCAAGCAGCTGATGGAACTGCTG 1647

Qy       1364 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1423
Db       1648 ACCAGAAAGGGATCCTTGCTGAACGAGAAAAACAAGGCTTTCCTCACGCCCCTGCATTTG 1707

Qy       1424 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1483
Db       1708 GCTGCCGAGCTGCTTCACTACGATGCCATGGAGGTGCTGCTAAAGCAGGGCGCCAAGGTT 1767

Qy       1484 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1543
Db       1768 AATGCATTGGACAGTCTTGGACAAACGCCACTGCATCGGTGCGC------CCGTGATGAG 1821

Qy       1544 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1603
Db       1822 CAAGCGGTGCGACTGCTGCTCTCGTACGCAGCGGACACGAATATCGTTTCCCTTGAGGGA 1881

Qy       1604 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1663
Db       1882 CTTACGGCCGCTCAATTGGCCTCGGACAGCGTGCTGAAGCTGCTCAAGAA---------- 1931

Qy       1664 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1723
Db       1932 -----TCCTCCGGACAGTGAGACACATTTACTGGAGGCAGCCAAGGCGGGAGATCTGGAC 1986

Qy       1724 ACTGT------AAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGG 1777
Db       1987 ACTGTGCGCCGTATAGTGCTCAACAATCCGATTTCGGTCAATTGCCGGGATTTGGACGGA 2046

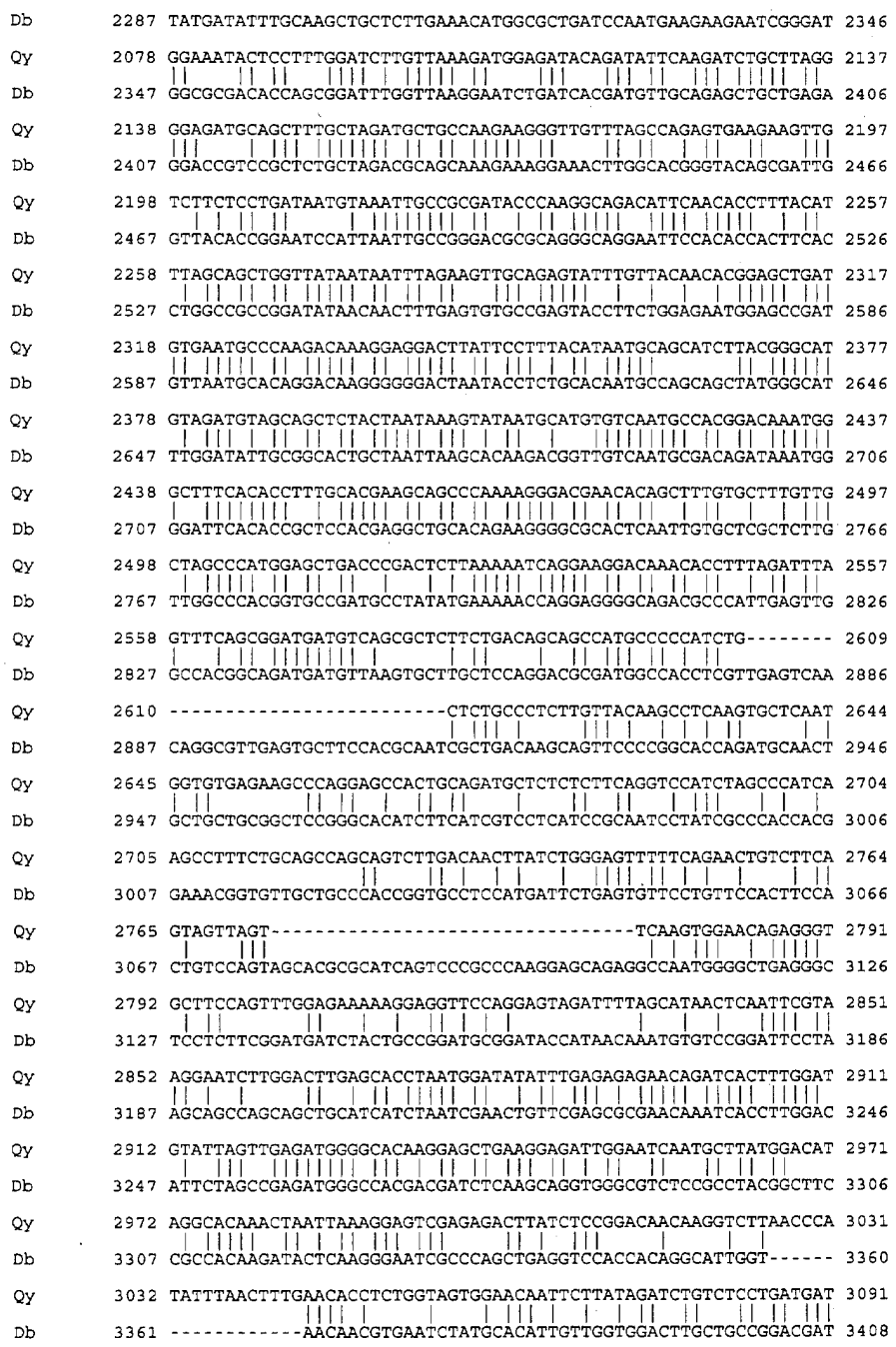
Qy       1778 CGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATAT 1837
Db       2047 CGACATTCCACACCTTTGCACTTTGCTGCTGGGTTTAATAGAGTGCCAGTGGTTCAGTTT 2106

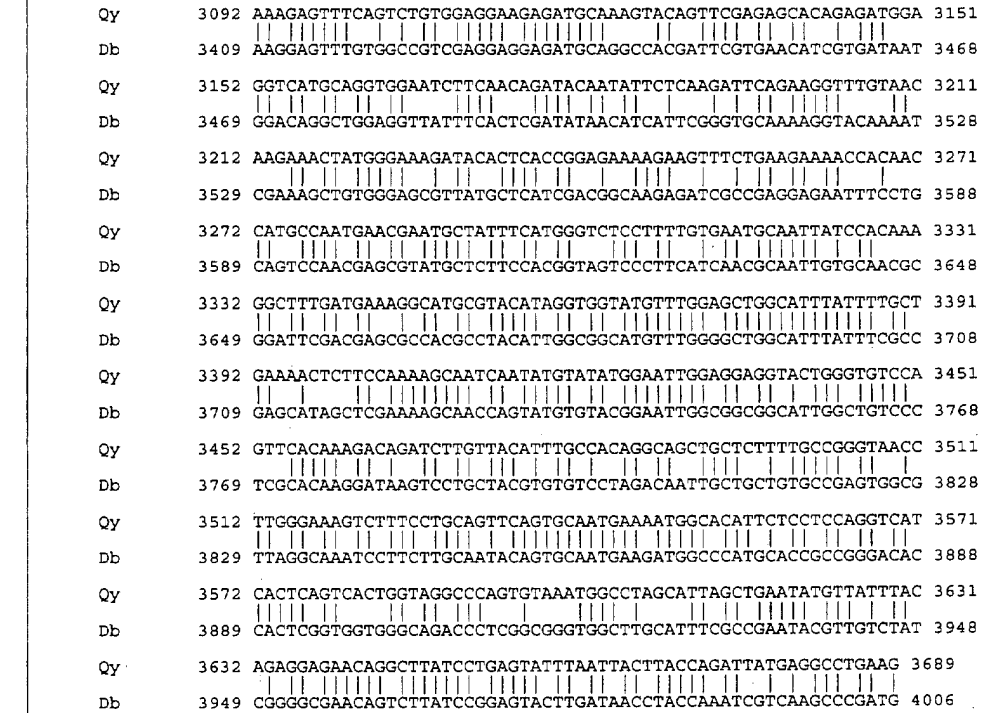
Qy       1838 CTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCAC 1897
Db       2107 CTTTTGGAACACGGCGCCGAGGTTTATGCGGCTGACAAGGGCGGACTGGTGCCCCTGCAC 2166

Qy       1898 AATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTA 1957
Db       2167 AATGCCTGCTCTTATGGGCACTATGAGGTAACCGAACTGCTGGTCAAGCACGGAGCCAAT 2226

Qy       1958 GTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAA 2017
Db       2227 GTAAATGTATCGGATTTGTGGAAGTTTACTCCTCTTCATGAAGCTGCCGCCAAGGGAAAG 2286

Qy       2018 TATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGAT 2077
```





```
RESULT 41
AAK99406
ID   AAK99406 standard; DNA; 2409 BP.
XX
AC   AAK99406;
XX
DT   27-JUN-2002  (first entry)
XX
DE   DNA of APP related human homologue hCP33787.
XX
KW   Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW   amyloid precursor protein; tissue-specific expression control; human APP;
KW   APP pathway modulator; gene therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .2409
FT                   /*tag=  a
FT                   /product= "Protein of human homologue hCP33787"
XX
PN   WO200226820-A2.
XX
PD   04-APR-2002.
XX
PF   01-OCT-2001; 2001WO-EP011345.
XX
PR   29-SEP-2000; 2000US-0236893P.
PR   14-JUN-2001; 2001US-0298309P.
XX
PA   (NOVS ) NOVARTIS AG.
```

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
PI Reinhardt MWHM, Zusman S;
XX
DR WPI; 2002-315796/35.
DR P-PSDB; AAO20512.
XX
PT New transgenic fly, containing DNA encoding an Abeta portion of human
PT APP, useful for identifying agents which modulate the APP pathway and
PT which can be used to treat Alzheimer's disease.
XX
PS Example 4; Page 112-113; 129pp; English.
XX
CC The invention relates to a transgenic fly whose genome comprises DNA
CC encoding a polypeptide having the Abeta portion of human amyloid
CC precursor protein (APP), fused to a signal sequence. The DNA sequence
CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC the specification. The DNA sequence is operably linked to a tissue-
CC specific expression control sequence. Expression of the sequence gives
CC the fly an altered phenotype. The purpose of the invention is for
CC identifying agents that inhibit or promote the expression and/or function
CC of genes or encoded polypeptides which modify the APP pathway. The agent
CC is a compound, triple helix DNA, antisense oligonucleotide, double
CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC to treat conditions such as Alzheimer's disease. The agent can be used as
CC an APP pathway modulator or in gene therapy. This polynucleotide sequence
CC represents the DNA of the APP related human homologue hCP33787
XX
SQ Sequence 2409 BP; 714 A; 480 C; 593 G; 622 T; 0 U; 0 Other;

Query Match 26.2%; Score 994; DB 6; Length 2409;
Best Local Similarity 65.4%; Pred. No. 8.3e-265;
Matches 2174; Conservative 0; Mismatches 20; Indels 1128; Gaps 4;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATAT----------------AAGAAAGATGA 744
              |||||||||||||||||||||||||| | | |               | |||   ||
Db        498 ATCTGCCAAAGCAGTGCTTACTGGTAAGTCTGTATACTCTGGTTATTCCAGGAAGCCTGT 557

Qy        745 ACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATT 804
              |      |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 AAAGAACAACCTTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATT 617

Qy        805 AAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGG 864
              ||||||||||||||||||||||||||||||
Db        618 AAATGTCAACTGCCACGCAAGTGATGGCAGA---------------------------- 648

Qy        865 ATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAA 924

Db        649 ------------------------------------------------------------ 648

Qy        925 AGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAAC 984
```

```
Db        649 ------------------------------------------------------------ 648

Qy        985 TGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCC 1044
                          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 ------------AAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCC 695

Qy       1045 TCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGG 1104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 TCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGG 755

Qy       1105 TGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACC 1164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        756 TGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACC 815

Qy       1165 ACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACG 1224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        816 ACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACG 875

Qy       1225 AGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCA 1284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        876 AGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCA 935

Qy       1285 TCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAA 1344
              ||||||||||||||||||||||
Db        936 TCCTCAAACACATGAAACAGCATTG---------------------------------- 960

Qy       1345 GCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATT 1404

Db        961 ------------------------------------------------------------ 960

Qy       1405 CTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGT 1464

Db        961 ------------------------------------------------------------ 960

Qy       1465 GAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGC 1524

Db        961 ------------------------------------------------------------ 960

Qy       1525 TGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAA 1584

Db        961 ------------------------------------------------------------ 960

Qy       1585 CATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACT 1644

Db        961 ------------------------------------------------------------ 960

Qy       1645 CCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGC 1704

Db        961 ------------------------------------------------------------ 960

Qy       1705 AAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAG 1764
                                    ||||||||||||||||||||||||||||||||||||||
Db        961 ----------------------------AAACTGTGTACTGTTCAGAGTGTCAACTGCAG 992

Qy       1765 AGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTC 1824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        993 AGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTC 1052

Qy       1825 CGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCT 1884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1053 CGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCT 1112

Qy       1885 TGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAA 1944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1113 TGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAA 1172

Qy       1945 ACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGC 2004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1173 ACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGC 1232

Qy       2005 AGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAA 2064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1233 AGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAA 1292
```

```
Qy       2065 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 2124
Db       1293 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 1352

Qy       2125 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 2184
Db       1353 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 1412

Qy       2185 AGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTC 2244
Db       1413 AGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTC 1472

Qy       2245 AACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACA 2304
Db       1473 AACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACA 1532

Qy       2305 ACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGC 2364
Db       1533 ACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGC 1592

Qy       2365 ATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGC 2424
Db       1593 ATCTTACGG---------------------------------------------------- 1601

Qy       2425 CACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCT 2484
Db       1602 ---------------------------------------------------------- 1601

Qy       2485 TTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAAC 2544
Db       1602 ---------------------------------------------------------- 1601

Qy       2545 ACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCC 2604
Db       1602 ---------------------------------------------------------- 1601

Qy       2605 ATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGC 2664
Db       1602 ---------------------------------------------------------- 1601

Qy       2665 CACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAG 2724
Db       1602 ---------------------------------------------------------- 1601

Qy       2725 TCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAC 2784
Db       1602 ---------------------------------------------------------- 1601

Qy       2785 AGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCA 2844
Db       1602 ---------------------------------------------------------- 1601

Qy       2845 ATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCAC 2904
Db       1602 ----------------------------------------------------GATCAC 1607

Qy       2905 TTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2964
Db       1608 TTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 1667

Qy       2965 TGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCT 3024
Db       1668 TGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCT 1727

Qy       3025 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 3084
Db       1728 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 1787

Qy       3085 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 3144
Db       1788 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 1847

Qy       3145 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 3204
Db       1848 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 1907

Qy       3205 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 3264
Db       1908 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 1967

Qy       3265 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3324
Db       1968 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 2027

Qy       3325 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 3384
Db       2028 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 2087

Qy       3385 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 3444
Db       2088 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 2147

Qy       3445 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 3504
Db       2148 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 2207

Qy       3505 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 3564
Db       2208 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 2267

Qy       3565 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 3624
Db       2268 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 2327

Qy       3625 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 3684
Db       2328 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 2387

Qy       3685 TGAAGGTATGGTCGATGGATAA 3706
Db       2388 TGAAGGTATGGTCGATGGATAA 2409
```

RESULT 42
AAC89703
ID AAC89703 standard; cDNA; 1007 BP.
XX
AC AAC89703;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human adult T-cell leukaemia cDNA clone MO-BC-203(5').
XX
KW Human; cytostatic; gene therapy; vaccine; breast cancer;
KW T cell leukaemia; cancer associated antigen; ss.
XX
OS Homo sapiens.
XX
PN WO200072021-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000WO-US014391.
XX
PR 26-MAY-1999; 99US-00320092.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Kuimov A, Lagarkova M, Koroleva E, Turetskaya R, Vdovichenko K;
PI Mescheryakov A, Litchinitser M, Kuprash D, Nedospasov S, Tureci O;
PI Sahin U, Pfruendschuh M, Old LJ, Knuth A, Jager E;
XX
DR WPI; 2001-032067/04.
DR P-PSDB; AAB48572.
XX

PT Diagnosing disorder characterized by expression of human cancer
PT associated antigen precursor, involves detecting interaction of an agent
PT with novel NA Group 1 nucleic acid molecule encoding the antigen
PT precursor.
XX
PS Example 1; Page 78; 87pp; English.
XX
CC The present sequence is given in a specification relating to nucleic
CC acids and encoded polypeptides which are cancer associated antigens
CC expressed in patients afflicted with breast cancer and/or T cell
CC leukaemia. A disorder characterised by expression of human cancer
CC associated antigen may be diagnosed by contacting a biological sample
CC with an agent that specifically binds to the nucleic acid that encodes
CC the antigen, complexed with human leukocyte antigen molecule, and
CC determining the interaction between the agent and the nucleic acid.
CC Cancer associated antigens, the nucleotides encoding them, antibodies
CC against them and the pharmaceutical compositions comprising them are
CC useful for diagnosing, monitoring and treating the diseases characterised
CC by the expression of one or more cancer associated antigens
XX
SQ Sequence 1007 BP; 315 A; 200 C; 219 G; 273 T; 0 U; 0 Other;

Query Match 26.1%; Score 990; DB 4; Length 1007;
Best Local Similarity 99.0%; Pred. No. 6.6e-264;
Matches 996; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy       978 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 1037
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAAT 60

Qy      1038 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1097
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 120

Qy      1098 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1157
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 180

Qy      1158 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1217
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240

Qy      1218 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1277
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 300

Qy      1278 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1337
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 360

Qy      1338 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1397
            ||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db       361 AAAGAAAGCAAATATGTGAACTGTTGCTAAAAAAAGGATCAAACATCAATGAAAAGACTA 420

Qy      1398 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1457
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480

Qy      1458 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1517
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540

Qy      1518 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1577
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 600

Qy      1578 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1637
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 660

Qy      1638 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1697
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 720

Qy      1698 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1757
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db       721 AAGCTGCAAAGGCTGGAGATGTCAAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 780

Qy      1758 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 840

Qy      1818 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1877
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       841 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGAGAAAG 900

Qy      1878 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 960

Qy      1938 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATT 1983
            |||||||||||||||||||  |||||  | || ||||||||||||
Db       961 TTGTTAAACATGGAGCAGTATTCAATGTTACCGAATTATGGAAATT 1006
```

RESULT 43
ADE53994
ID ADE53994 standard; cDNA; 797 BP.
XX
AC ADE53994;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human prostate cancer cDNA #341.
XX
KW Human; prostate cancer; ss; cDNA combination; differential expression;
KW gene.
XX
OS Homo sapiens.
XX
PN US2003190640-A1.
XX
PD 09-OCT-2003.
XX
PF 29-MAY-2002; 2002US-00252157.
XX
PR 31-MAY-2001; 2001US-0295048P.
XX
PA (FARI/) FARIS M.
PA (PEAR/) PEARSON C I.
XX
PI Faris M, Pearson CI;
XX
DR WPI; 2003-831619/77.
XX
PT New combination comprising cDNAs that are differentially expressed in
PT prostate cancer, useful for diagnosing, treating or monitoring the
PT progression of treatment of prostate cancer.
XX
PS Claim 1; SEQ ID NO 341; 42pp; English.
XX
CC The invention relates to a combination comprising a number of cDNAs
CC expressed in prostate cancer. The invention also relates to a method for
CC detecting differential expression of one or more cDNAs in a sample
CC containing nucleic acids by hybridising a substrate with the nucleic
CC acids, thus forming one or more hybridisation complexes, detecting
CC hybridisation complex formation and comparing the complexes formed with
CC standard complexes, where differences between the standard and the sample
CC complex formation indicate differential expression of cDNAs in the
CC sample. The differential expression is diagnostic of prostate cancer. The
CC invention also relates to proteins and antibodies related to the cDNAs.
CC The combination is useful for diagnosing, treating or monitoring the
CC progression of treatment of prostate cancer. The antibodies are useful
CC for detecting prostate cancer. This sequence represents a human prostate
CC cancer cDNA of the invention.
XX

```
SQ   Sequence 797 BP; 256 A; 144 C; 184 G; 212 T; 0 U; 1 Other;

  Query Match           20.9%;  Score 794.4;  DB 9;  Length 797;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-209;
  Matches  795;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1344 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 60

Qy       1404 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 120

Qy       1464 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 180

Qy       1524 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 240

Qy       1584 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 1643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACATTATATCCCTTCANGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 300

Qy       1644 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 360

Qy       1704 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 420

Qy       1764 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1823
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 480

Qy       1824 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 540

Qy       1884 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 600

Qy       1944 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 660

Qy       2004 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2063
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        661 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 720

Qy       2064 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 780

Qy       2124 AAGATCTGCTTAGGGGA 2140
             ||||||||||||||||
Db        781 AAGATCTGCTTAGGGGA 797


RESULT 44
AAC77578
ID   AAC77578 standard; cDNA; 1344 BP.
XX
AC   AAC77578;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF3133 polynucleotide sequence SEQ ID NO:6265.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US008621.
XX
PR   31-MAR-1999;   99US-0127607P.
PR   02-APR-1999;   99US-0127636P.
PR   05-APR-1999;   99US-0127728P.
PR   30-MAR-2000; 2000US-00540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB43369.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease.
XX
PS   Claim 5; Page 5454; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC   anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC   cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC   dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC   antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC   sequences can be used for determining the presence of or predisposition
CC   to, or preventing or treating pathological conditions associated with an
CC   ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC   proteins in gene therapy vectors. The proteins and nucleic acids may be
CC   used to treat cancers, proliferative disorders, neurodegenerative
CC   disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC   diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC   storage, systemic lupus erythematosus, severe combined immunodeficiency
CC   (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC   disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC   cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC   enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ   Sequence 1344 BP; 418 A; 243 C; 259 G; 422 T; 0 U; 2 Other;

  Query Match           20.7%;  Score 786;  DB 3;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-207;
  Matches  786;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3009 CCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTC 3068
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 CCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTC 85

Qy       3069 TTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTA 3128
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        86 TTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTA 145

Qy      3129 CAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTC 3188
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       146 CAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTC 205

Qy      3189 TCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAG 3248
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       206 TCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAG 265

Qy      3249 AAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTT 3308
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       266 AAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTT 325

Qy      3309 TTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGT 3368
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       326 TTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGT 385

Qy      3369 TTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAA 3428
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       386 TTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAA 445

Qy      3429 TTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGC 3488
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       446 TTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGC 505

Qy      3489 AGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAA 3548
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       506 AGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAA 565

Qy      3549 TGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAG 3608
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       566 TGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAG 625

Qy      3609 CATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTT 3668
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       626 CATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTT 685

Qy      3669 ACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAAT 3728
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       686 ACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAAT 745

Qy      3729 TCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAA 3788
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       746 TCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAA 805

Qy      3789 AAAAAA 3794
                ||||||
Db       806 AAAAAA 811
```

RESULT 45
AAH08231
ID AAH08231 standard; cDNA; 838 BP.
XX
AC AAH08231;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:5066.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 1; SEQ ID NO 5066; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 838 BP; 260 A; 143 C; 203 G; 227 T; 0 U; 5 Other;

Query Match 20.7%; Score 784.2; DB 4; Length 838;
Best Local Similarity 98.0%; Pred. No. 8.6e-207;
Matches 823; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

```
Qy      2715 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 2774
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 CAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTT 60

Qy      2775 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 2834
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTA 120

Qy      2835 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 2894
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAG 180

Qy      2895 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 2954
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAA 240

Qy      2955 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 3014
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGAC 300

Qy      3015 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 3074
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAG 360

Qy      3075 ATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 3134
```

```
Db        361 ATCTGTCTCCTGATGATAAAGAGTTTCCGTCTGTGGAGGAAGAGATGCAAAGTACAGTTC 420

Qy       3135 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 3194
Db        421 GAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGA 480

Qy       3195 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3254
Db        481 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 540

Qy       3255 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3314
Db        541 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 600

Qy       3315 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3374
Db        601 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 660

Qy       3375 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3434
Db        661 CTGGCATTTATTTTGCTGAAAACTCTT-CAAAAGCAATCAATATGTATATGGAATTGGAN 719

Qy       3435 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3494
Db        720 GANGT-CTGGGTGTCCAGTTCACAAAGACCGATCTTGGTACATTTGCCACANGCAGCTGC 778

Qy       3495 TCTTTTG--CCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 3552
Db        779 TCTTTTGGCCGGGTAACCCTTGGGAAAGTCTTTTCTGCAGTTCANTGCAATGNAAATGGC 838

Search completed: March 29, 2004, 20:37:43
Job time : 1434.1 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 19:26:40 ; Search time 14330.6 Seconds
(without alignments)
11484.026 Million cell updates/sec

Title: US-09-843-159B-1
Perfect score: 3797
Sequence: 1 ctttgaagacactggatttc.....cctttgctgaaaaaaaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3797	100.0	3797	6	AR392872	AR392872 Sequence
2	3393.4	89.4	3816	6	AR392873	AR392873 Sequence
3	3388.8	89.2	4275	6	AR361486	AR361486 Sequence
4	3388.8	89.2	6028	9	AF309033	AF309033 Homo sapi
5	3388.8	89.2	6123	6	AX057579	AX057579 Sequence
6	3387.2	89.2	3815	9	AF329696	AF329696 Homo sapi
7	3387.2	89.2	4127	6	AX062241	AX062241 Sequence
8	3387.2	89.2	4406	6	AX062247	AX062247 Sequence
9	3387.2	89.2	4992	6	AX062275	AX062275 Sequence
10	3387.2	89.2	5002	6	AX062273	AX062273 Sequence
11	3387.2	89.2	5810	9	AF264912	AF264912 Homo sapi
12	3387.2	89.2	6189	9	AF438201	AF438201 Homo sapi
13	3386	89.2	3394	6	AR352002	AR352002 Sequence
14	3386	89.2	3394	6	AX133694	AX133694 Sequence
15	3382.8	89.1	4512	6	AR232161	AR232161 Sequence
16	3382.8	89.1	4512	6	AX076213	AX076213 Sequence
17	3338.4	87.9	4295	9	AF305081	AF305081 Homo sapi
18	3300.4	86.9	3501	9	AF342982	AF342982 Homo sapi
19	3297	86.8	3400	6	AX029397	AX029397 Sequence
20	3297	86.8	3400	6	BD075726	BD075726 Potential
21	3296.2	86.8	3508	6	AX062142	AX062142 Sequence
22	3291.8	86.7	3498	6	AR232162	AR232162 Sequence
23	3291.8	86.7	3498	6	AX076214	AX076214 Sequence
24	2969.6	78.2	4493	6	AR361484	AR361484 Sequence
25	2968.2	78.2	4297	6	AR361485	AR361485 Sequence
26	2774.6	73.1	5005	6	AX052530	AX052530 Sequence
27	2439.8	64.3	2971	6	AX062233	AX062233 Sequence
28	2439.8	64.3	3353	6	AX062229	AX062229 Sequence
29	2403.2	63.3	3930	5	AY142107	AY142107 Gallus ga
30	2357.8	62.1	3799	6	AX062231	AX062231 Sequence
31	1922.4	50.6	4130	9	HSM804989	AL833676 Homo sapi
32	1839.8	48.5	4054	10	BC057370	BC057370 Mus muscu
33	1813.4	47.8	3984	6	AX062144	AX062144 Sequence
34	1813.4	47.8	4134	6	BD235110	BD235110 Tankyrase
35	1813.4	47.8	4134	6	AR274521	AR274521 Sequence
36	1813.4	47.8	4134	6	AR361487	AR361487 Sequence
37	1813.4	47.8	4134	6	AX467061	AX467061 Sequence
38	1813.4	47.8	4134	9	AF082556	AF082556 Homo sapi
39	1808.8	47.6	3804	5	AY142108	AY142108 Gallus ga
40	1794.8	47.3	3254	6	AX057580	AX057580 Sequence
41	1794.8	47.3	8901	9	AF082557	AF082557 Homo sapi
42	1449.6	38.2	4491	6	BD235115	BD235115 Tankyrase
43	1449.6	38.2	4491	6	AR274526	AR274526 Sequence
44	1449.6	38.2	4491	6	AX467067	AX467067 Sequence
45	1359.4	35.8	1568	6	AX062239	AX062239 Sequence

ALIGNMENTS

RESULT 1
AR392872
LOCUS AR392872 3797 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6617102.
ACCESSION AR392872
VERSION AR392872.1 GI:40118097
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 3797)
AUTHORS Luo,Y., Chan,E., Xu,X., Huang,B. and Ossovskaya,V.
TITLE Tankyrase H, compositions involved in the cell cycle and methods of use
JOURNAL Patent: US 6617102-A 1 09-SEP-2003;

```
FEATURES             Location/Qualifiers
     source          1. .3797
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             100.0%;  Score 3797;  DB 6;  Length 3797;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTTTGAAGACACTGGATTTCATACTTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTTGAAGACACTGGATTTCATACTTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60

Qy         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120

Qy        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180

Qy        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240

Qy        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300

Qy        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360

Qy        361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420

Qy        421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480

Qy        481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540

Qy        541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600

Qy        601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660

Qy        661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720

Qy        721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780

Qy        781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840

Qy        841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900

Qy        901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960

Qy        961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020

Qy       1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080

Qy       1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140

Qy       1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200

Qy       1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260

Qy       1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320

Qy       1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380

Qy       1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440

Qy       1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500

Qy       1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560

Qy       1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620

Qy       1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680

Qy       1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740

Qy       1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800

Qy       1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860

Qy       1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920

Qy       1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980

Qy       1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
```

```
Qy       2041 CCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Db       2041 CCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100

Qy       2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Db       2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160

Qy       2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Db       2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220

Qy       2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Db       2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280

Qy       2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Db       2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340

Qy       2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Db       2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400

Qy       2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Db       2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460

Qy       2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520
Db       2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520

Qy       2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580
Db       2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580

Qy       2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640
Db       2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640

Qy       2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Db       2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700

Qy       2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Db       2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760

Qy       2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820
Db       2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820

Qy       2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880
Db       2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880

Qy       2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940
Db       2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940

Qy       2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000
Db       2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000

Qy       3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Db       3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060

Qy       3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120
Db       3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120

Qy       3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180
Db       3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180

Qy       3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240
Db       3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240

Qy       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
Db       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300

Qy       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
Db       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360

Qy       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Db       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420

Qy       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
Db       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480

Qy       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
Db       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540

Qy       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
Db       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600

Qy       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660
Db       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660

Qy       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720
Db       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720

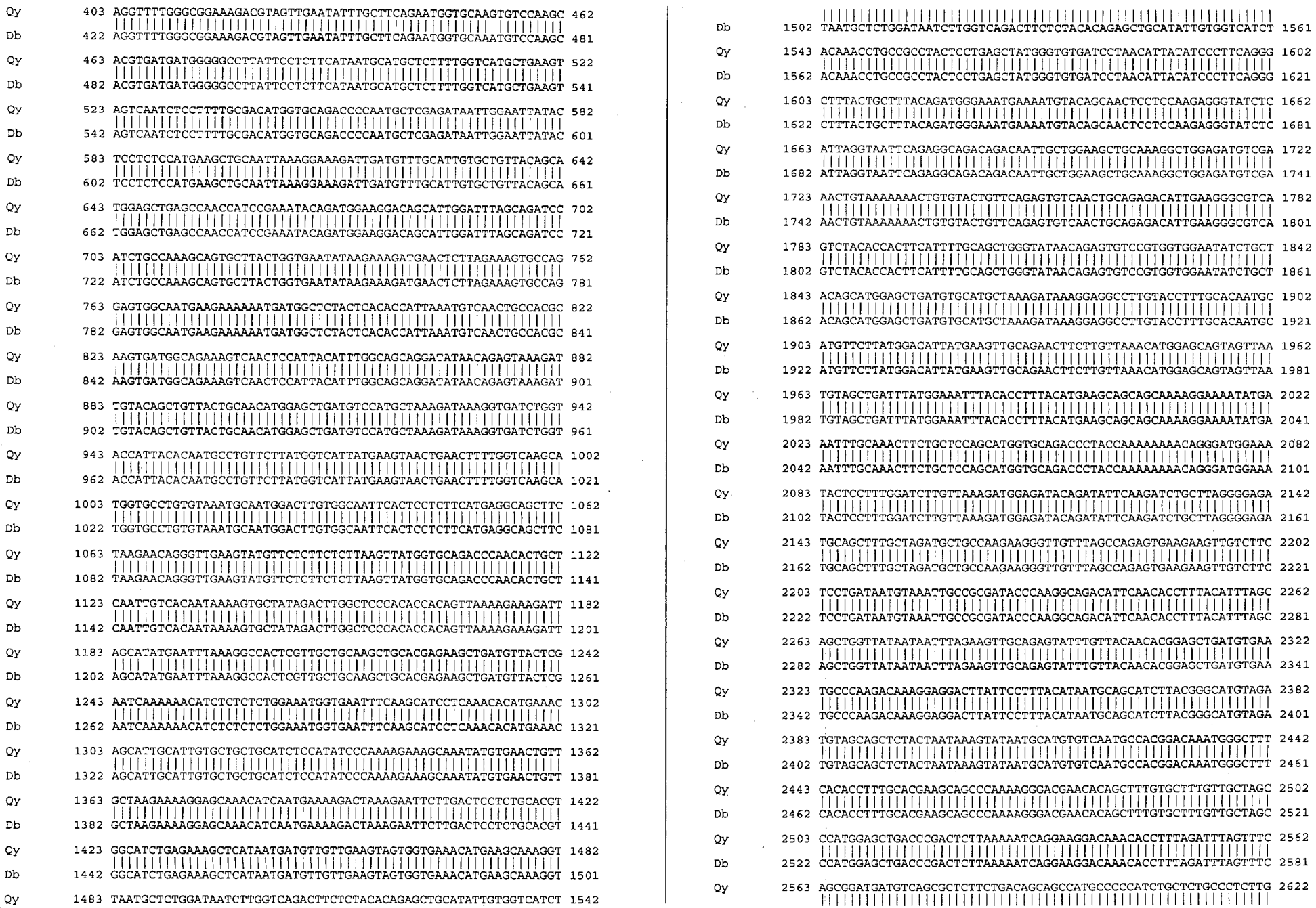
Qy       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780
Db       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780

Qy       3781 TTGCTGAAAAAAAAAAA 3797
Db       3781 TTGCTGAAAAAAAAAAA 3797
```

RESULT 2
AR392873

```
LOCUS       AR392873               3816 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 2 from patent US 6617102.
ACCESSION   AR392873
VERSION     AR392873.1  GI:40118098
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3816)
  AUTHORS   Luo,Y., Chan,E., Xu,X., Huang,B. and Ossovskaya,V.
  TITLE     Tankyrase H, compositions involved in the cell cycle and methods of
            use
  JOURNAL   Patent: US 6617102-A 2 09-SEP-2003;
FEATURES             Location/Qualifiers
     source          1. .3816
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match 89.4%; Score 3393.4; DB 6; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


Qy 403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db 422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481
Qy 463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db 482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541
Qy 523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db 542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601
Qy 583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db 602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661
Qy 643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db 662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721
Qy 703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db 722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781
Qy 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db 782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841
Qy 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db 902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2101
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622

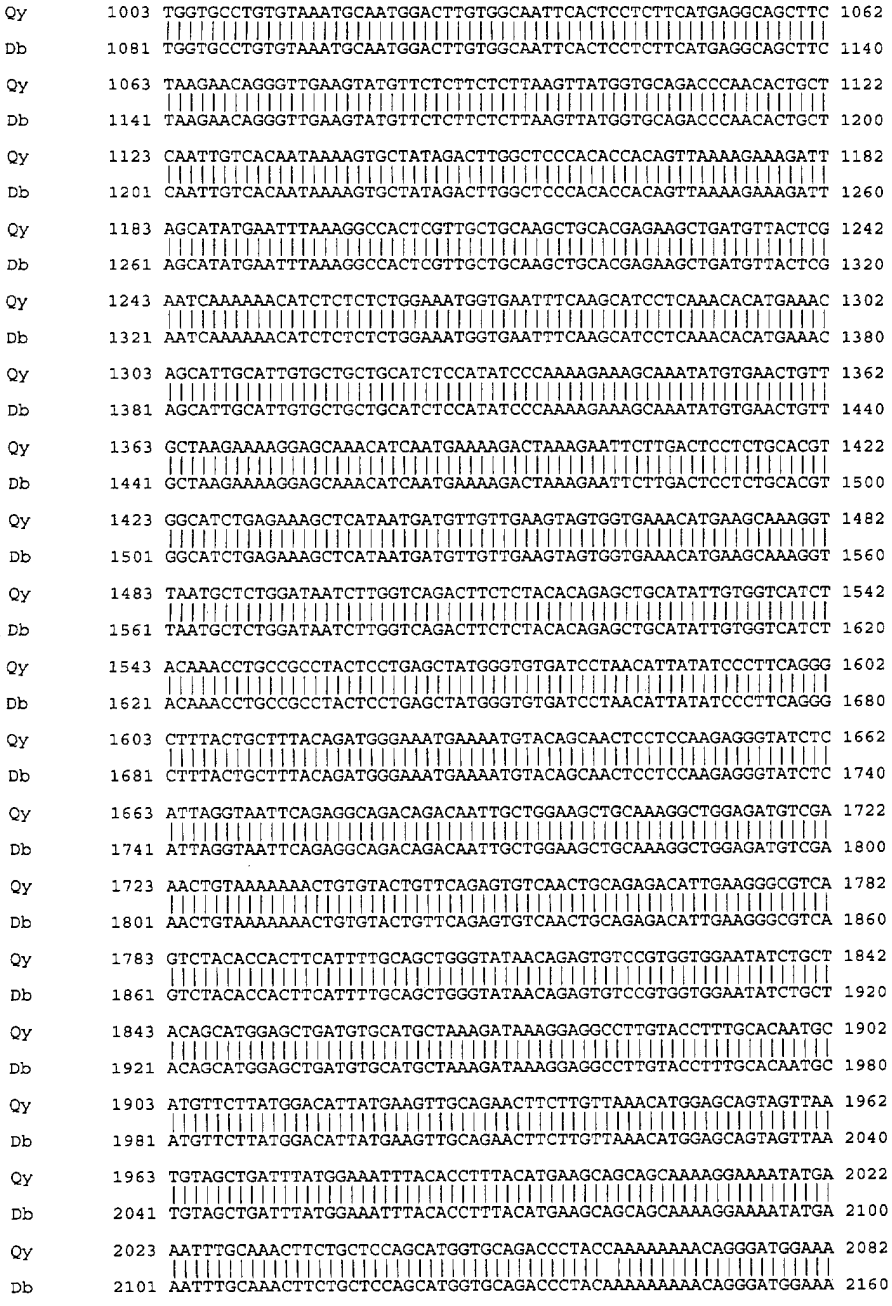
```
Db       2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721
```

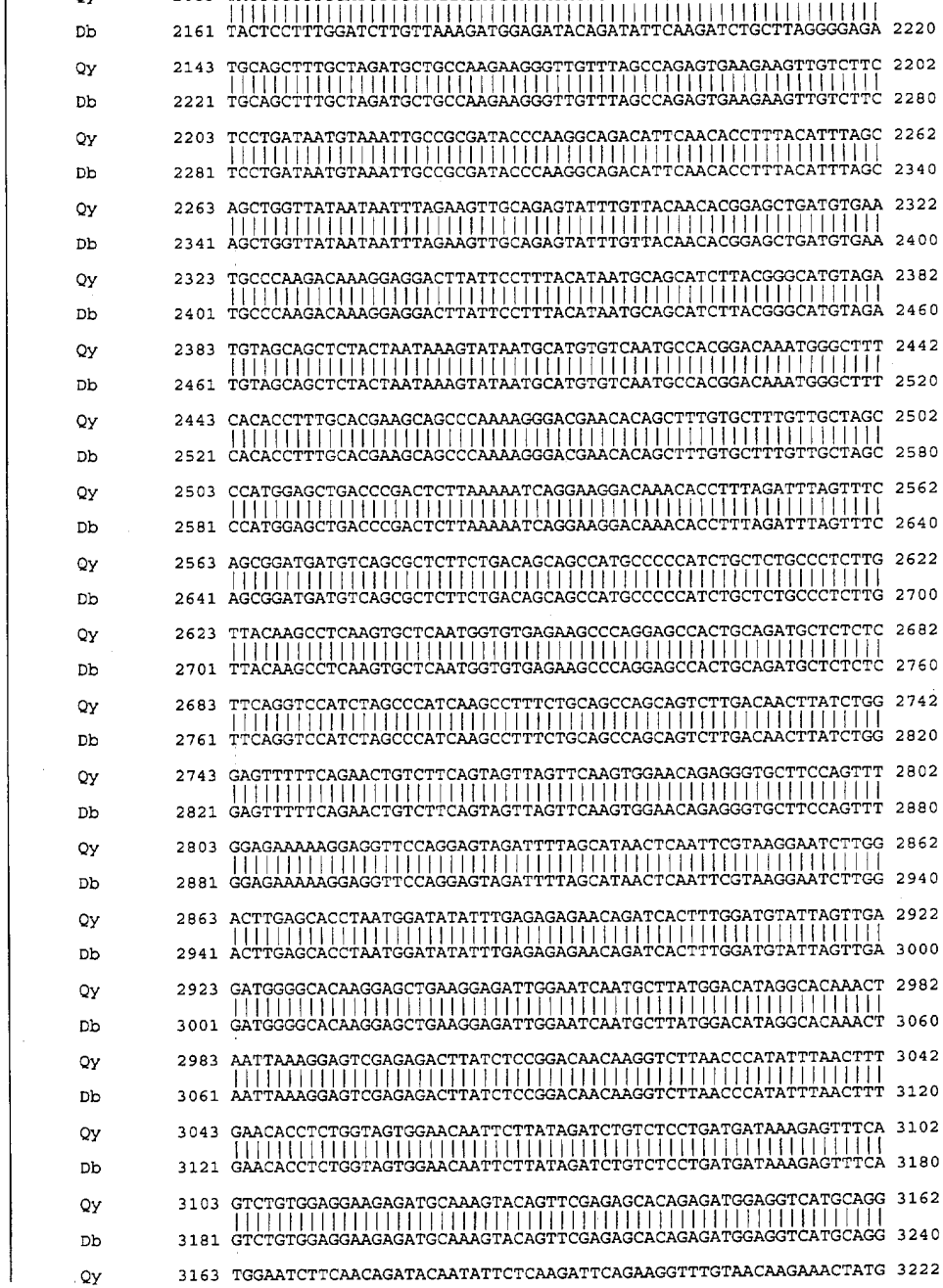
```
Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
             |||||||||||||||||||||||||||||||||||
Db       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816


RESULT 3
AR361486
LOCUS       AR361486               4275 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 5 from patent US 6599728.
ACCESSION   AR361486
VERSION     AR361486.1  GI:33769334
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4275)
  AUTHORS   Morin,G.B., Funk,W.D. and Piatyszek,M.A.
  TITLE     Second mammalian tankyrase
  JOURNAL   Patent: US 6599728-A 5 29-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .4275
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           89.2%;  Score 3388.8;  DB 6;  Length 4275;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3390;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 540
Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 600
Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 660
Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 720
Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 780
Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 840
Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 900
Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 960
Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1020
Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1080
```


Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 1081 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1140
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1141 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1200
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1201 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1260
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1261 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1320
Qy 1243 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1321 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1380
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1381 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1440
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1441 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1500
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1501 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1560
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1561 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1620
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1621 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1680
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1681 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1740
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1741 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1800
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1801 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1860
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1861 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1920
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1921 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1980
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1981 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2040
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2041 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2100
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 2101 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2160


Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2161 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2220
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2221 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2280
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2281 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2340
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2341 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2400
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2401 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2460
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2461 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2520
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2521 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2580
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2581 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2640
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2641 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2700
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2701 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2760
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2761 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2820
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2821 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2880
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2881 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2940
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2941 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3000
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 3001 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3060
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3061 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3120
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3121 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3180
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3181 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3240
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

```
Db       3241 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3300

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

Db       3301 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3360

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342

Db       3361 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3420

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402

Db       3421 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3480

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462

Db       3481 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3540

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

Db       3541 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3600

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Db       3601 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3660

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Db       3661 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3720

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Db       3721 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3780

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Db       3781 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3840

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794

Db       3841 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

```
RESULT 4
AF309033
LOCUS       AF309033               6028 bp    mRNA    linear   PRI 31-JUL-2001
DEFINITION  Homo sapiens tankyrase-2 (TNKS-2) mRNA, complete cds.
ACCESSION   AF309033
VERSION     AF309033.1  GI:15042551
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6028)
  AUTHORS   Chi,N.-W. and Lodish,H.F.
  TITLE     A novel homolog of Tankyrase
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 6028)
  AUTHORS   Chi,N.-W. and Lodish,H.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-2000) Whitehead Institute for Biomedical
            Research, 9 Cambridge Center, Cambridge, MA 02142, USA
FEATURES             Location/Qualifiers
     source          1. .6028
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q23.1"
                     /note="STS WI-16054"
     gene            1. .6028
                     /gene="TNKS-2"
     CDS             252. .3752
                     /gene="TNKS-2"
                     /note="poly(ADP-ribose) polymerase containing ankyrin
                     repeats and SAM (sterile alpha module) motif"
                     /codon_start=1
                     /product="tankyrase-2"
                     /protein_id="AAK82330.1"
                     /db_xref="GI:15042552"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
     polyA_signal    6010. .6015
                     /gene="TNKS-2"
ORIGIN
```

```
  Query Match             89.2%;  Score 3388.8;  DB 9;  Length 6028;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3390;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Db        449 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 508

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522

Db        509 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 568

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582

Db        569 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 628

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642

Db        629 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 688

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702

Db        689 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 748

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762

Db        749 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 808

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822

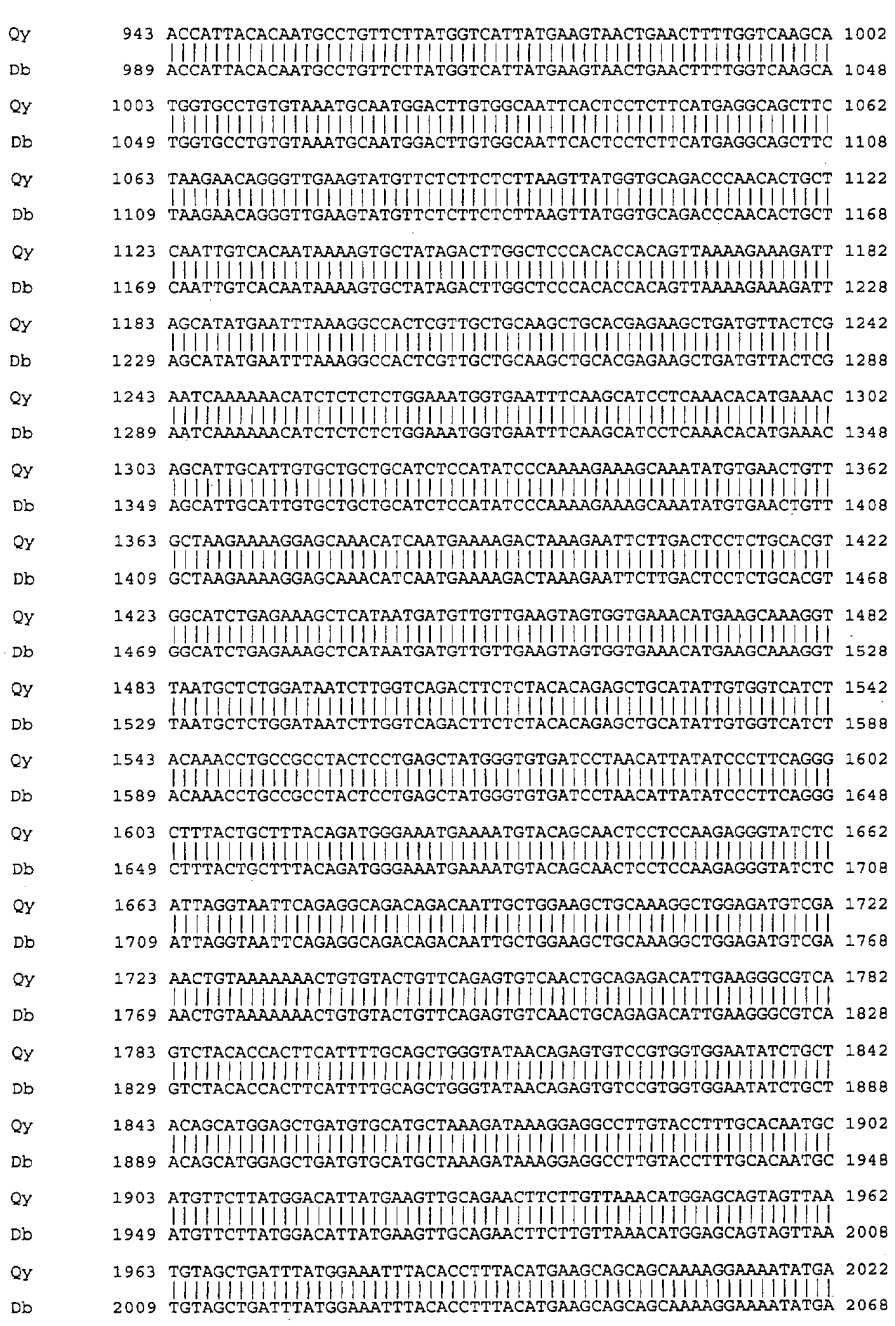
Db        809 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 868

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882

Db        869 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 928

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942

Db        929 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 988
```


Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 989 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1048
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 1049 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1108
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1109 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1168
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1169 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1228
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1229 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1288
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1289 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1348
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1349 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1408
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1409 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1468
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1469 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1528
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1529 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1588
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1589 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1648
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1649 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1708
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1709 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1768
Qy 1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1769 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1828
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1829 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1888
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1889 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1948
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1949 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2008
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2009 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2068
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2069 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2128
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2129 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2188
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2189 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2248
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2249 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2308
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2309 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2368
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2369 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2428
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2429 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2488
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2489 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2548
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2549 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2608
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2609 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2668
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2669 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2728
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2729 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2788
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2789 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2848
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2849 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2908
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2909 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2968
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2969 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3028
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3029 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3088
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3089 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3148
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162

```
Db       3149 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3208

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3209 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3268

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3269 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3328

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3329 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3388

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3389 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3448

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3449 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3508

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3509 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3568

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3569 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3628

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3629 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3688

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3748

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3808

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3809 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3840


RESULT 5
AX057579
LOCUS       AX057579                6123 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 2 from Patent WO0077225.
ACCESSION   AX057579
VERSION     AX057579.1  GI:12310301
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chi,N.W. and Lodish,H.F.
  TITLE     A novel insulin signaling molecule
  JOURNAL   Patent: WO 0077225-A 2 21-DEC-2000;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; THE GENERAL
            HOSPITAL CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .6123
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           89.2%;  Score 3388.8;  DB 6;  Length 6123;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3390;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        449 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 508

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 568

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 628

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 688

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 748

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 808

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        809 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 868

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        869 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 928

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        929 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 988

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        989 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1048

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1049 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1108

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1109 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1168

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1169 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1228

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1229 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1288

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1289 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1348

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1349 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1408

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1409 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1468

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1469 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1528
Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db      1529 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1588
Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db      1589 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1648
Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db      1649 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1708
Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db      1709 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1768
Qy      1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db      1769 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1828
Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db      1829 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1888
Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db      1889 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1948
Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db      1949 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2008
Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db      2009 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2068
Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db      2069 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2128
Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db      2129 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2188
Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db      2189 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2248
Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db      2249 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2308
Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db      2309 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2368
Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db      2369 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2428
Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db      2429 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2488
Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db      2489 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2548
Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db      2549 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2608
```

```
Qy      2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db      2609 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2668
Qy      2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db      2669 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2728
Qy      2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db      2729 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2788
Qy      2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db      2789 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2848
Qy      2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db      2849 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2908
Qy      2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db      2909 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2968
Qy      2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db      2969 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3028
Qy      2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db      3029 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3088
Qy      3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db      3089 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3148
Qy      3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db      3149 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3208
Qy      3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db      3209 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3268
Qy      3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db      3269 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3328
Qy      3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db      3329 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3388
Qy      3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db      3389 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3448
Qy      3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db      3449 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3508
Qy      3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db      3509 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3568
Qy      3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db      3569 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3628
Qy      3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db      3629 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3688
```

```
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3689 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3748

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3749 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3808

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3809 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3840


RESULT 6
AF329696
LOCUS       AF329696               3815 bp    mRNA    linear   PRI 14-MAY-2001
DEFINITION  Homo sapiens tankyrase 2 mRNA, complete cds.
ACCESSION   AF329696
VERSION     AF329696.1  GI:13161041
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3815)
  AUTHORS   Lyons,R.J., Deane,R., Lynch,D.K., Ye,Z.S., Sanderson,G.M.,
            Eyre,H.J., Sutherland,G.R. and Daly,R.J.
  TITLE     Identification of a novel human tankyrase through its interaction
            with the adaptor protein Grb14
  JOURNAL   J. Biol. Chem. 276 (20), 17172-17180 (2001)
  MEDLINE   21264473
   PUBMED   11278563
REFERENCE   2  (bases 1 to 3815)
  AUTHORS   Lyons,R.J., Deane,R., Lynch,D.K., Ye,Z.-S.J., Sanderson,G.M.,
            Eyre,H.J., Sutherland,G.R. and Daly,R.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-DEC-2000) Cancer Research, Garvan Institute of
            Medical Research, 384 Victoria St., Sydney, NSW 2010, Australia
FEATURES             Location/Qualifiers
     source          1. .3815
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q23.2"
     CDS             140. .3640
                     /codon_start=1
                     /product="tankyrase 2"
                     /protein_id="AAK13463.1"
                     /db_xref="GI:13161042"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN
```

```
Query Match           89.2%;  Score 3387.2;  DB 9;  Length 3815;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        337 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 396

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 456

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 516

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 576

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 636

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 696

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 756

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 816

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 876

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 936

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 996

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1056

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1116

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1176

Qy       1243 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1236

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1296

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1356

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
```

```
Db      1357 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1416
Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db      1417 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1476
Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db      1477 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1536
Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db      1537 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1596
Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db      1597 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1656
Qy      1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db      1657 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1716
Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db      1717 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1776
Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db      1777 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1836
Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db      1837 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1896
Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db      1897 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1956
Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db      1957 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2016
Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db      2017 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2076
Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db      2077 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2136
Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db      2137 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2196
Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db      2197 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2256
Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db      2257 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2316
Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db      2317 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2376
Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db      2377 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2436
Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
```

```
Db      2437 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2496
Qy      2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db      2497 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2556
Qy      2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db      2557 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2616
Qy      2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db      2617 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2676
Qy      2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db      2677 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2736
Qy      2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db      2737 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2796
Qy      2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db      2797 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2856
Qy      2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db      2857 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2916
Qy      2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db      2917 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2976
Qy      3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db      2977 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3036
Qy      3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db      3037 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3096
Qy      3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db      3097 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3156
Qy      3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db      3157 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3216
Qy      3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db      3217 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3276
Qy      3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db      3277 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3336
Qy      3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db      3337 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3396
Qy      3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db      3397 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3456
Qy      3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db      3457 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3516
Qy      3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db      3517 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3576
```

```
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3577 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3636

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3637 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3696

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3697 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3728


RESULT 7
AX062241
LOCUS       AX062241               4127 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 100 from Patent WO0100849.
ACCESSION   AX062241
VERSION     AX062241.1  GI:12540142
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 100 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .4127
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .3511
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25060.1"
                     /db_xref="GI:12540143"
                     /db_xref="REMTREMBL:CAC25060"
                     /translation="ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVVK
                     RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNAC
                     SFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG
                     RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPL
                     HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM
                     DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFK
                     GHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK
                     GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ
                     TCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV
                     ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPL
                     HNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKK
                     NRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRH
                     STPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNAC
                     VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLT
                     AAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSS
                     VVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKE
                     LKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE
                     EEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER
                     MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK
                     DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYR
                     GEQAYPEYLITYQIMRPEGMVDG"
     3'UTR           3509. .4127
ORIGIN

  Query Match          89.2%;  Score 3387.2;  DB 6;  Length 4127;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

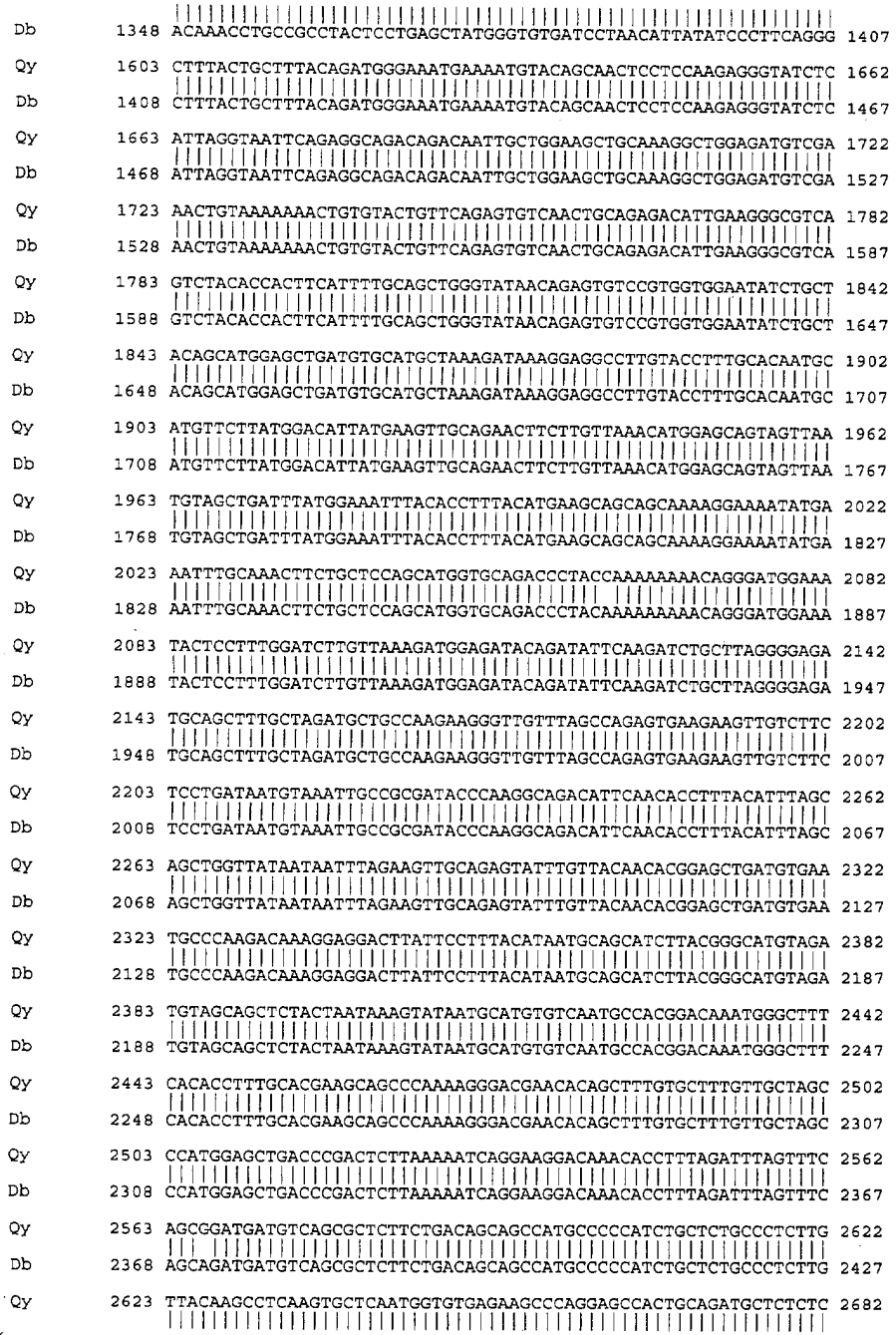
Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

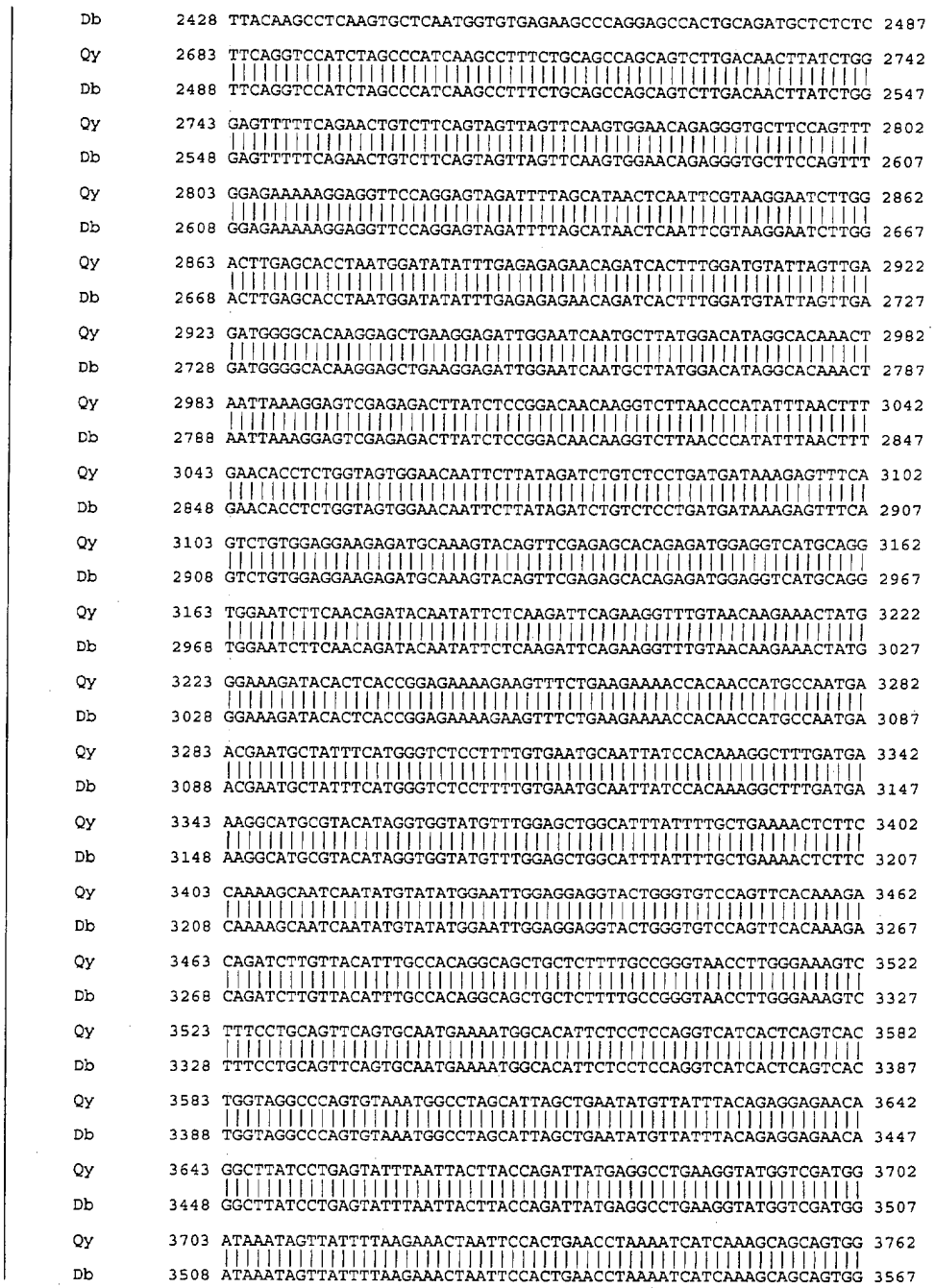
Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
```


Db 1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1587
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1588 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1647
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1648 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1707
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1708 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1767
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1768 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1827
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1887
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1888 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1947
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1948 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2007
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2008 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2067
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2068 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2127
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2188 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2247
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2248 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2307
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2308 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2427
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682


Db 2428 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2487
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2488 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2547
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2607
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2608 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2728 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2787
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2847
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 2848 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2907
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 2908 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2967
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 2968 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3027
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3028 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3087
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3088 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3207
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3208 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3267
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3268 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3327
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3328 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3387
Qy 3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db 3388 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3447
Qy 3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db 3448 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3507
Qy 3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db 3508 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3567

```
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||||
Db       3568 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3599


RESULT 8
AX062247
LOCUS       AX062247               4406 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 106 from Patent WO0100849.
ACCESSION   AX062247
VERSION     AX062247.1  GI:12540148
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 106 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .4406
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .3790
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25062.1"
                     /db_xref="GI:12540149"
                     /db_xref="REMTREMBL:CAC25062"
                     /translation="ELAGGALPASAAASFQDPDGGFALPPPPRGSRGAGSPARGARGR
                     GHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMSGRRC
                     AGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKSTPLHF
                     AAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPNARDN
                     WNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEYKKDE
                     LLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHGADVH
                     AKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEVCSLL
                     LSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRIKKHLSLE
                     MVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASEKAH
                     NDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTAL
                     QMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQST
                     PLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVN
                     VADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
                     GDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHG
                     ADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQL
                     CALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSP
                     GATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDF
                     SITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLI
                     SGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRY
                     NILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAY
                     IGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFL
                     QFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG
                     "
ORIGIN

  Query Match           89.2%;  Score 3387.2;  DB 6;  Length 4406;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 546

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 606

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 666


Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 726

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 786

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 846

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 906

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 966

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1026

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1086

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1146

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1206

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1266

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1326

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1386

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1446

Qy       1363 GCTAAGAAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 GCTAAGAAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1506

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1566

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1626

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1686

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1746
```

```
Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1806

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1866

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1926

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1986

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2046

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2106

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       2107 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2166

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2226

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2286

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2346

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2347 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2406

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2407 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2466

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2526

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2527 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2586

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2587 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2646

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2647 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2706

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2707 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2766

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2767 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2826

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2886

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2887 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2946

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2947 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3006

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3007 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3066

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3067 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3126

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3127 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3186

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3187 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3246

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3247 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3306

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3307 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3366

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3367 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3426

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3427 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3486

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3487 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3546

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3547 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3606

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3607 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3666

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3667 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3726

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3727 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3786

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3787 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3846

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3847 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3878

RESULT 9
```

```
AX062275
LOCUS       AX062275               4992 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 134 from Patent WO0100849.
ACCESSION   AX062275
VERSION     AX062275.1  GI:12540176
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 134 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .4992
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             876. .4376
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25067.1"
                     /db_xref="GI:12540177"
                     /db_xref="REMTREMBL:CAC25067"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match           89.2%;  Score 3387.2;  DB 6;  Length 4992;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db       1073 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 1132

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 1192

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 1252

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 1312

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 1372

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1432

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1492

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1552

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1612

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1672

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1732

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1792

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1852

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1912

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1972

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 2032

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 2092

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 2152

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 2212

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 2272

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 2332

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 2392

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2452

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
```

```
Db      2453 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2512
Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db      2513 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 2572
Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db      2573 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2632
Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db      2633 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2692
Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db      2693 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2752
Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db      2753 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2812
Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db      2813 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2872
Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db      2873 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2932
Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db      2933 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2992
Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db      2993 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 3052
Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db      3053 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 3112
Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db      3113 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 3172
Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db      3173 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 3232
Qy      2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db      3233 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 3292
Qy      2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db      3293 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 3352
Qy      2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db      3353 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 3412
Qy      2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db      3413 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 3472
Qy      2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db      3473 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3532
Qy      2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db      3533 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3592
Qy      2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db      3593 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3652
Qy      2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db      3653 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3712
Qy      3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db      3713 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3772
Qy      3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db      3773 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3832
Qy      3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db      3833 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3892
Qy      3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db      3893 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3952
Qy      3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db      3953 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 4012
Qy      3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db      4013 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 4072
Qy      3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db      4073 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 4132
Qy      3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db      4133 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 4192
Qy      3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db      4193 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 4252
Qy      3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db      4253 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 4312
Qy      3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db      4313 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 4372
Qy      3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db      4373 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 4432
Qy      3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db      4433 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4464
```

```
RESULT 10
AX062273
LOCUS       AX062273               5002 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 132 from Patent WO0100849.
ACCESSION   AX062273
VERSION     AX062273.1  GI:12540174
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 132 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .5002
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             229. .4386
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25066.1"
                     /db_xref="GI:12540175"
                     /db_xref="REMTREMBL:CAC25066"
                     /translation="MWKLKTRTISTVRLPKECARSEQNFQGLFFSVMVVRKFSLESPG
                     LKTTTTKNTICRIVRLQQNPPQRWRWDEAPSPAAEASRLTFPTNPSRRLASRNLPSRC
                     PATAHARDDVTCAPGAGRSWQELAGGALPASAAASFQDPDGGFALPPPPRGSRGAGSP
                     ARGARGRGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAAR
                     IMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGR
                     KSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGA
                     DPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLT
                     GEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLL
                     QHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNR
                     VEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHSLLQAAREADVTRI
                     KKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLH
                     VASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIIS
                     LQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRD
                     IEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLV
                     KHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDT
                     DIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVA
                     EYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAA
                     QKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQV
                     LNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKK
                     EVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLI
                     KGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHA
                     GGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKG
                     FDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRV
                     TLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMR
                     PEGMVDG"
ORIGIN

  Query Match           89.2%;  Score 3387.2;  DB 6;  Length 5002;
  Best Local Similarity 99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 1142

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 1202

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 1262

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 1322

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 1382

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1442

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1502

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1562

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1563 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1622

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1623 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1682

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1683 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1742

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1743 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1802

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1862

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1922

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1982

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 2042

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2043 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 2102

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2103 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 2162

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2163 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 2222

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 2282

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2283 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 2342

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2343 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 2402

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2403 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2462

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2463 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2522

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2523 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 2582
Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       2583 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2642
Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       2643 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2702
Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db       2703 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2762
Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       2763 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2822
Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       2823 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2882
Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2883 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2942
Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2943 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 3002
Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       3003 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 3062
Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       3063 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 3122
Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       3123 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 3182
Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       3183 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 3242
Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       3243 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 3302
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       3303 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 3362
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       3363 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 3422
Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       3423 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 3482
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       3483 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3542
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       3543 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3602
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       3603 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3662
```

```
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       3663 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3722
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       3723 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3782
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       3783 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3842
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       3843 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3902
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3903 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3962
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3963 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 4022
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       4023 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 4082
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       4083 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 4142
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       4143 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 4202
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       4203 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 4262
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       4263 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 4322
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       4323 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 4382
Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       4383 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 4442
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       4443 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4474
```

RESULT 11
AF264912
LOCUS AF264912 5810 bp mRNA linear PRI 14-MAY-2001
DEFINITION Homo sapiens tankyrase-like protein (TNKL) mRNA, complete cds.
ACCESSION AF264912
VERSION AF264912.1 GI:12005975
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5810)
AUTHORS Kuimov,A.N., Kuprash,D.V., Petrov,V.N., Vdovichenko,K.K.,
Scanlan,M.J., Jongeneel,C.V., Lagarkova,M.A. and Nedospasov,S.A.
TITLE Cloning and characterization of TNKL, a member of tankyrase gene

```
            family
  JOURNAL   Genes Immun. 2 (1), 52-55 (2001)
  MEDLINE   21190090
   PUBMED   11294570
REFERENCE   2  (bases 1 to 5810)
  AUTHORS   Kuimov,A.N., Kuprash,D.V., Petrov,V.N., Vdovichenko,K.K.,
            Kashkarova,U.A., Boitchenko,V.E., Nedospasov,S.A. and
            Lagarkova,M.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAY-2000) Laboratory of Molecular Immunology, A.N.
            Belozersky Institute of Physico-Chemical Biology, Moscow State
            University, Moscow 119899, Russia
FEATURES             Location/Qualifiers
     source          1. .5810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q22-10q24"
                     /tissue_type="breast carcinoma"
                     /note="derived from clones MO-BC-203 and I203"
     gene            <1. .5810
                     /gene="TNKL"
                     /note="a new member of tankyrase gene family; current
                     Unigene claster Hs.168491"
     CDS             1. .3501
                     /gene="TNKL"
                     /note="a new member of tankyrase family"
                     /codon_start=1
                     /product="tankyrase-like protein"
                     /protein_id="AAG44694.1"
                     /db_xref="GI:12005976"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
     STS             complement(1132. .1277)
                     /note="SHGC-53837"
     STS             complement(3665. .3975)
                     /note="WI-16054"
     STS             complement(5544. .5777)
                     /note="WI-8339"
ORIGIN

  Query Match             89.2%;  Score 3387.2;  DB 9;  Length 5810;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 557

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 617

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 677

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 737

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 797

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 857

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 917

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 977

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1037

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1097

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1098 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1157

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1158 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1217

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1218 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1277

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1278 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1337

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1338 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1397

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
```

```
Db      1398 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1457
Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db      1458 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1517
Qy      1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db      1518 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1577
Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db      1578 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1637
Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db      1638 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1697
Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db      1698 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1757
Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db      1758 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1817
Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db      1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1877
Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db      1878 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1937
Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db      1938 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 1997
Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db      1998 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2057
Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db      2058 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2117
Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db      2118 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2177
Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db      2178 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2237
Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db      2238 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2297
Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db      2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357
Qy      2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db      2358 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417
Qy      2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db      2418 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2477
Qy      2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db      2478 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2537
Qy      2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db      2538 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2597
Qy      2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db      2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2657
Qy      2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db      2658 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2717
Qy      2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db      2718 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2777
Qy      2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db      2778 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2837
Qy      3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db      2838 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2897
Qy      3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db      2898 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2957
Qy      3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db      2958 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3017
Qy      3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db      3018 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3077
Qy      3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db      3078 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3137
Qy      3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db      3138 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3197
Qy      3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db      3198 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3257
Qy      3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db      3258 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3317
Qy      3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db      3318 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3377
Qy      3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db      3378 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3437
Qy      3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db      3438 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3497
Qy      3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db      3498 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3557
Qy      3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db      3558 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3589
```

```
RESULT 12
AF438201
LOCUS       AF438201               6189 bp    mRNA    linear   PRI 12-DEC-2001
DEFINITION  Homo sapiens tankyrase II mRNA, complete cds.
ACCESSION   AF438201
VERSION     AF438201.1  GI:17530294
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6189)
  AUTHORS   Yin,Y. and Gelmann,E.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2001) Lombardi Cancer Center, Georgetown
            University, 3800 Reservoir Rd, NW, Washington, DC 20007, USA
FEATURES             Location/Qualifiers
     source          1. .6189
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q24"
     CDS             380. .3880
                     /note="telomerase binding protein"
                     /codon_start=1
                     /product="tankyrase II"
                     /protein_id="AAL40795.1"
                     /db_xref="GI:17530295"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match             89.2%;  Score 3387.2;  DB 9;  Length 6189;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 636

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 696

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 756

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 816

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 876

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 936

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 996

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1056

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1116

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1176

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1236

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1296

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1356

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1416

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1476

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1536

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1596

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1656

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1657 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1716

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1717 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1776

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1777 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1836

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1837 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1896
```

```
Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db       1897 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1956

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db       1957 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2016

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db       2017 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 2076

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       2077 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2136

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       2137 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2196

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db       2197 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2256

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       2257 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2316

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       2317 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2376

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2377 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2436

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2437 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2496

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2497 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2556

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2557 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2616

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2617 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2676

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2677 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2736

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2737 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2796

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2797 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2856

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2857 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2916

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2917 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2976

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862

Db       2977 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3036

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       3037 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3096

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       3097 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3156

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       3157 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3216

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       3217 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3276

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       3277 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3336

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       3337 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3396

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3397 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3456

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3457 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3516

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3517 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3576

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3577 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3636

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3637 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3696

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3697 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3756

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3757 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3816

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3817 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3876

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       3877 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3936

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       3937 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3968
```

RESULT 13
AR352002
LOCUS AR352002 3394 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 1 from patent US 6589725.
ACCESSION AR352002

```
VERSION     AR352002.1  GI:33756953
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3394)
  AUTHORS   Luo,Y., Chan,E., Xu,X. and Huang,B.
  TITLE     Tankyrase H, compositions involved in the cell cycle and methods of
            use
  JOURNAL   Patent: US 6589725-A 1 08-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .3394
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             89.2%;  Score 3386;  DB 6;  Length 3394;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        404 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCA 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          1 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 60

Qy        464 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 120

Qy        524 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 180

Qy        584 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 240

Qy        644 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 703
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAGCTGAGCCAACCATCCTAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 300

Qy        704 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 360

Qy        764 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 420

Qy        824 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480

Qy        884 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 540

Qy        944 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 600

Qy       1004 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 660

Qy       1064 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 720

Qy       1124 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 780

Qy       1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 840

Qy       1244 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 900

Qy       1304 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 960

Qy       1364 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1020

Qy       1424 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1080

Qy       1484 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1140

Qy       1544 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1200

Qy       1604 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1260

Qy       1664 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1320

Qy       1724 ACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1380

Qy       1784 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1440

Qy       1844 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1500

Qy       1904 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1560

Qy       1964 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620

Qy       2024 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 2083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 1680

Qy       2084 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCATTATCTGCTTAGGGGAGAT 1740

Qy       2144 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 1800

Qy       2204 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 1860

Qy       2264 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1861 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 1920
Qy       2324 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2383
Db       1921 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 1980
Qy       2384 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2443
Db       1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2040
Qy       2444 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2503
Db       2041 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2100
Qy       2504 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2563
Db       2101 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160
Qy       2564 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2623
Db       2161 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2220
Qy       2624 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2683
Db       2221 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2280
Qy       2684 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2743
Db       2281 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2340
Qy       2744 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2803
Db       2341 AGTTTTTCAGAACTGTCTTCATTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2400
Qy       2804 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2863
Db       2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2460
Qy       2864 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2923
Db       2461 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2520
Qy       2924 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2983
Db       2521 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2580
Qy       2984 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 3043
Db       2581 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 2640
Qy       3044 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 3103
Db       2641 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 2700
Qy       3104 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 3163
Db       2701 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 2760
Qy       3164 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3223
Db       2761 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 2820
Qy       3224 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3283
Db       2821 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 2880
Qy       3284 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3343
Db       2881 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 2940
Qy       3344 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3403
Db       2941 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3000
```

```
Qy       3404 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3463
Db       3001 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3060
Qy       3464 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3523
Db       3061 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3120
Qy       3524 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3583
Db       3121 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3180
Qy       3584 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643
Db       3181 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3240
Qy       3644 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
Db       3241 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3300
Qy       3704 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3763
Db       3301 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3360
Qy       3764 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
Db       3361 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3394
```

```
RESULT 14
AX133694
LOCUS       AX133694               3394 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 1 from Patent WO0130987.
ACCESSION   AX133694
VERSION     AX133694.1  GI:14139712
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Luo,Y., Chan,E., Xu,X. and Huang,B.
  TITLE     Tankyrase h, compositions involved in the cell cycle and methods of
            use
  JOURNAL   Patent: WO 0130987-A 1 03-MAY-2001;
            Rigel Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .3394
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             89.2%;  Score 3386;  DB 6;  Length 3394;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        404 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCA 463
Db          1 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 60
Qy        464 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 523
Db         61 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 120
Qy        524 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 583
Db        121 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 180
Qy        584 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 643
Db        181 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 240
```

```
Qy       644 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 703
Db       241 GGAGCTGAGCCAACCATCCTAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 300

Qy       704 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 763
Db       301 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 360

Qy       764 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 823
Db       361 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 420

Qy       824 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883
Db       421 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480

Qy       884 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 943
Db       481 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 540

Qy       944 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1003
Db       541 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 600

Qy      1004 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1063
Db       601 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 660

Qy      1064 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1123
Db       661 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 720

Qy      1124 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1183
Db       721 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 780

Qy      1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1243
Db       781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 840

Qy      1244 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1303
Db       841 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 900

Qy      1304 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1363
Db       901 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 960

Qy      1364 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1423
Db       961 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1020

Qy      1424 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1483
Db      1021 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1080

Qy      1484 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1543
Db      1081 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1140

Qy      1544 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1603
Db      1141 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1200

Qy      1604 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1663
Db      1201 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1260

Qy      1664 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1723
Db      1261 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1320

Qy      1724 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1783
Db      1321 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1380

Qy      1784 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1843
Db      1381 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1440

Qy      1844 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1903
Db      1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1500

Qy      1904 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1963
Db      1501 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1560

Qy      1964 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2023
Db      1561 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620

Qy      2024 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 2083
Db      1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAAT 1680

Qy      2084 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2143
Db      1681 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCATTATCTGCTTAGGGGAGAT 1740

Qy      2144 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2203
Db      1741 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 1800

Qy      2204 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2263
Db      1801 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 1860

Qy      2264 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2323
Db      1861 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 1920

Qy      2324 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2383
Db      1921 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 1980

Qy      2384 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2443
Db      1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2040

Qy      2444 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2503
Db      2041 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2100

Qy      2504 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2563
Db      2101 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160

Qy      2564 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2623
Db      2161 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2220

Qy      2624 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2683
Db      2221 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2280

Qy      2684 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2743
Db      2281 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2340

Qy      2744 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2803
Db      2341 AGTTTTTCAGAACTGTCTTCATTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2400

Qy      2804 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2863
```

```
Db       2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2460

Qy       2864 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2923
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2520

Qy       2924 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2983
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2580

Qy       2984 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 3043
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 2640

Qy       3044 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 3103
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 2700

Qy       3104 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 3163
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 2760

Qy       3164 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3223
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 2820

Qy       3224 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3283
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 2880

Qy       3284 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 2940

Qy       3344 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3000

Qy       3404 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3060

Qy       3464 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3120

Qy       3524 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3583
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3180

Qy       3584 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3240

Qy       3644 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3300

Qy       3704 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3360

Qy       3764 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
             ||||||||||||||||||||||||||||||||||
Db       3361 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3394


RESULT 15
AR232161
LOCUS       AR232161               4512 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 3 from patent US 6455290.
ACCESSION   AR232161
VERSION     AR232161.1  GI:27274048
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4512)
  AUTHORS   Berthelsen,J., Toma,S. and Isacchi,A.
  TITLE     Tankyrase homolog protein (THP), nucleic acids, and methods related
            to the same
  JOURNAL   Patent: US 6455290-A 3 24-SEP-2002;
FEATURES             Location/Qualifiers
     source          1. .4512
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             89.1%;  Score 3382.8;  DB 6;  Length 4512;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3381;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 279

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 339

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 399

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 459

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 519

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 579

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 639

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 699

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 759

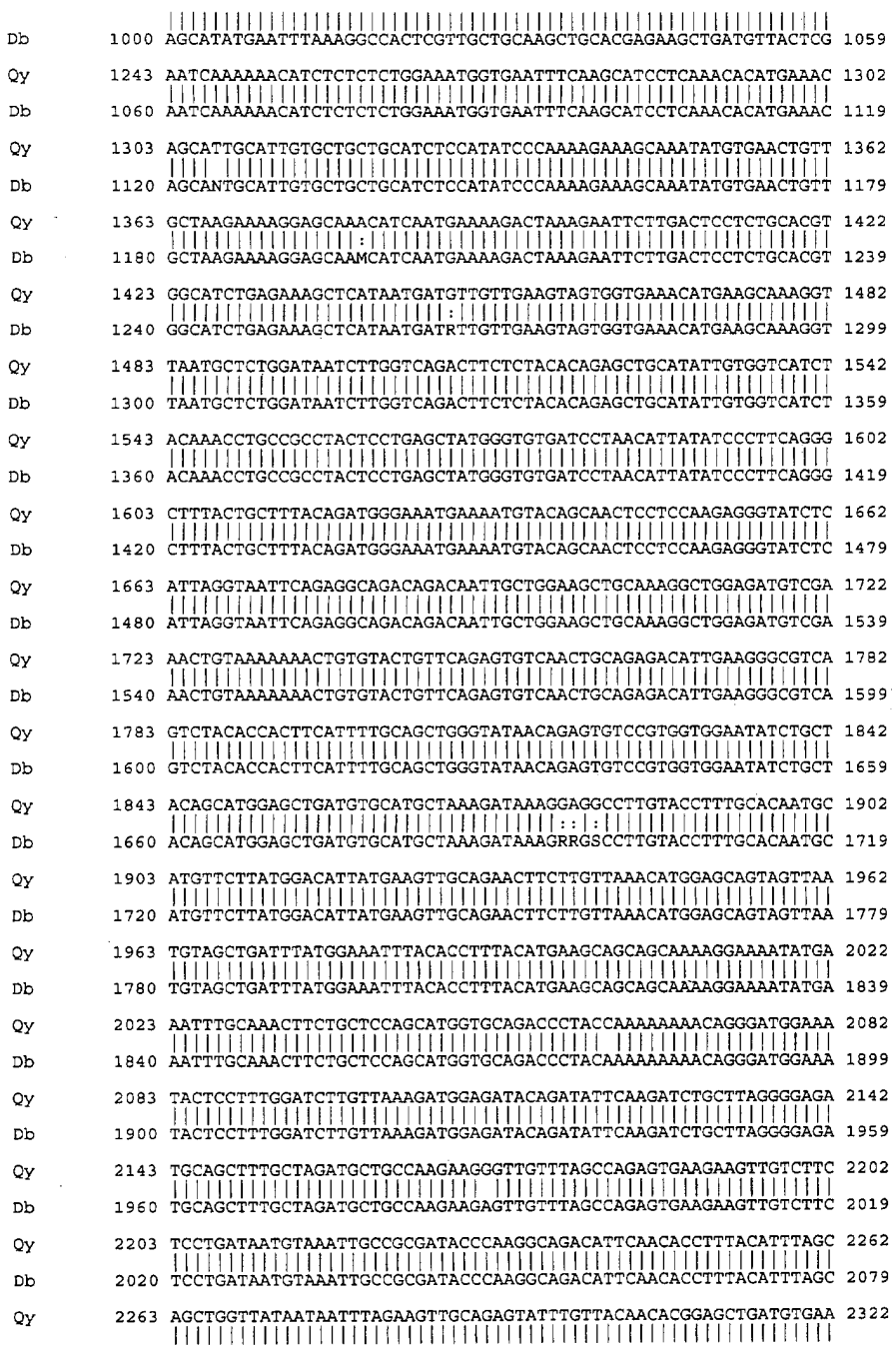
Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 819

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 879

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 939

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 999

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
```


Db 1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1059
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1060 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1119
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1120 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1179
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1180 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1239
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1240 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1299
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1300 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1359
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1360 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1419
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1420 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1479
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1480 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1539
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1540 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1599
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1600 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1659
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1660 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1719
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1720 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1779
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1780 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1839
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 1840 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1899
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1900 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1959
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1960 TGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2019
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2020 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2079
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2080 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2139
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2140 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2199
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2200 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2259
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2260 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGKTGCTAGC 2319
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2320 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2379
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2380 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2439
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2440 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2499
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2500 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2559
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2560 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2619
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2620 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGNATCTTGG 2679
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2680 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2739
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2740 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2799
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2800 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2859
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 2860 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2919
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 2920 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2979
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 2980 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3039
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3040 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3099
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3100 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3159
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3160 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3219

```
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3220 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3279

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3280 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3339

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3340 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3399

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3400 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3459

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3460 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3519

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3520 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3579

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3580 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3611
```

```
RESULT 16
AX076213
LOCUS       AX076213               4512 bp    DNA     linear   PAT 06-FEB-2001
DEFINITION  Sequence 3 from Patent WO0104326.
ACCESSION   AX076213
VERSION     AX076213.1  GI:12710838
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Berthelsen,J., Toma,S. and Isacchi,A.
  TITLE     Tankyrase homolog protein (thp), nucleic acids, and methods related
            to the same
  JOURNAL   Patent: WO 0104326-A 3 18-JAN-2001;
            PHARMACIA & UPJOHN S.p.A. (IT)
FEATURES             Location/Qualifiers
     source          1. .4512
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     unsure          1124
                     /note="n is any base"
     unsure          2672
                     /note="n is any base"
ORIGIN

  Query Match           89.1%;  Score 3382.8;  DB 6;  Length 4512;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3381;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        220 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 279

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 339

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 399

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 459

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 519

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 579

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 639

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 699

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 759

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 819

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 879

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 939

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 999

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1059

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1119

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1179

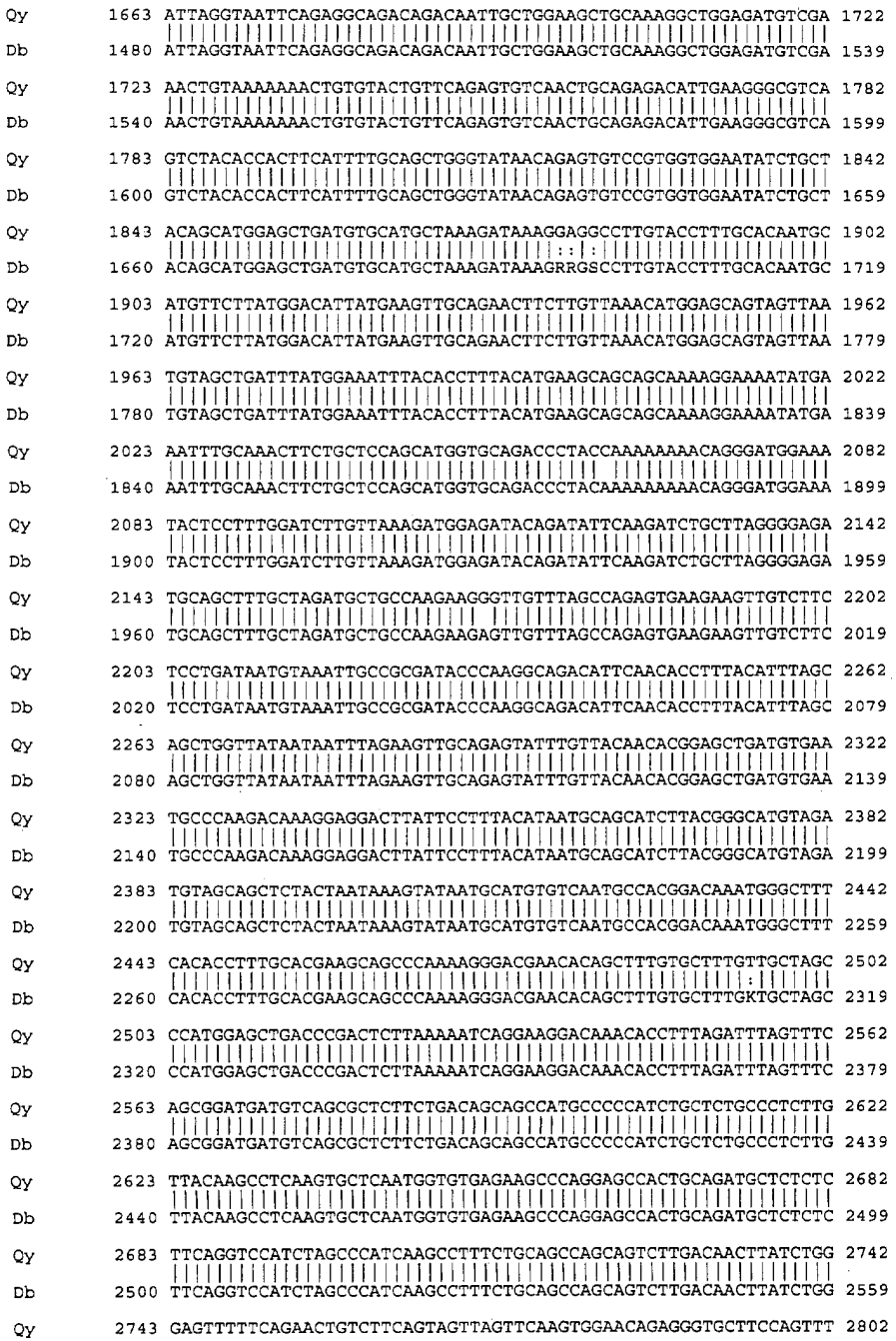
Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1239

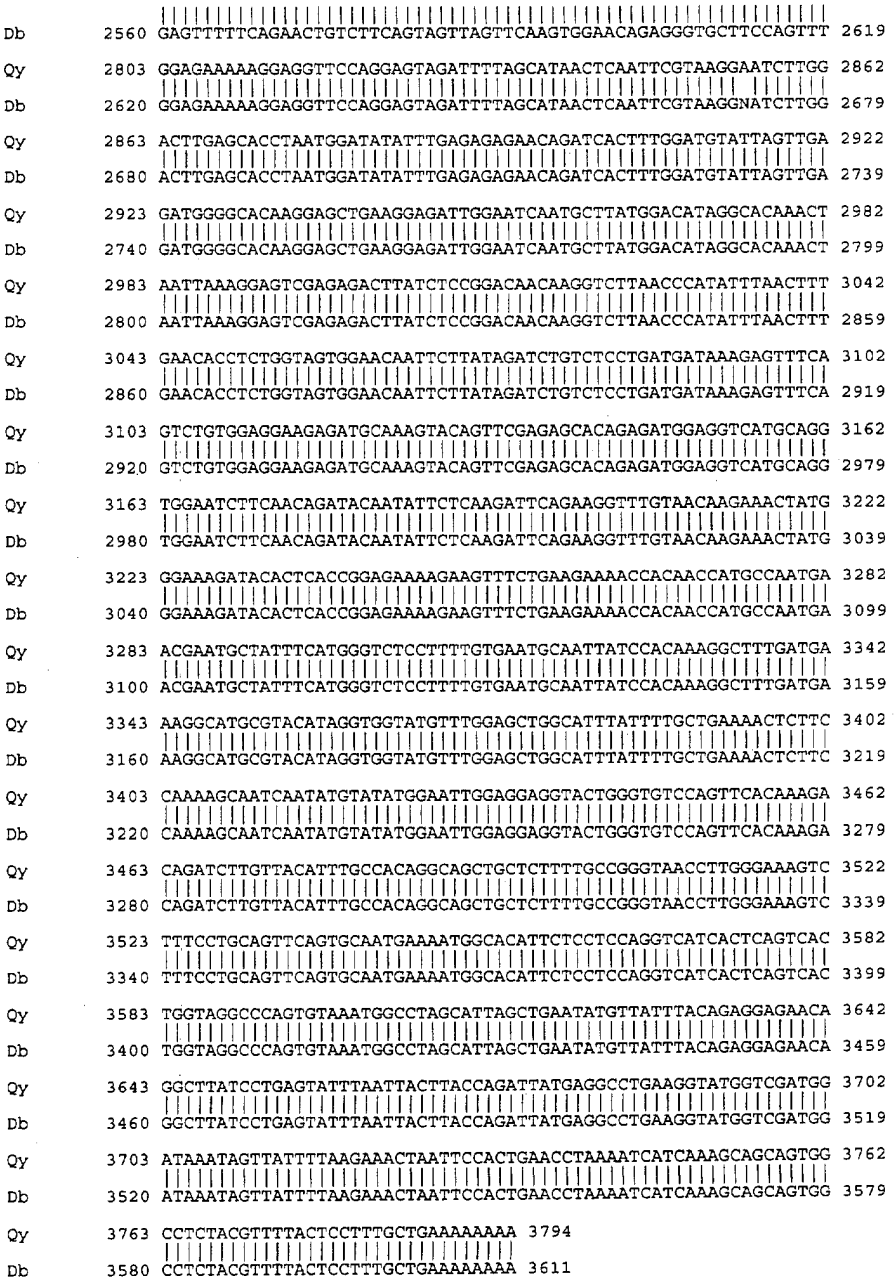
Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db       1240 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1299

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1359

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1360 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1419

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1420 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1479
```





RESULT 17

```
AF305081
LOCUS       AF305081                4295 bp    mRNA    linear   PRI 17-APR-2001
DEFINITION  Homo sapiens tankyrase-related protein mRNA, partial cds.
ACCESSION   AF305081
VERSION     AF305081.1  GI:10953951
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4295)
  AUTHORS   Monz,D., Munnia,A., Comtesse,N., Fischer,U., Steudel,W.I.,
            Feiden,W., Glass,B. and Meese,E.U.
  TITLE     Novel tankyrase-related gene detected with meningioma-specific sera
  JOURNAL   Clin. Cancer Res. 7 (1), 113-119 (2001)
  MEDLINE   21072550
   PUBMED   11205898
REFERENCE   2  (bases 1 to 4295)
  AUTHORS   Monz,D.W. and Meese,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Human Genetics, Universitaet des
            Saarlandes, Geb. 60, Homburg 66421, Germany
FEATURES             Location/Qualifiers
     source          1. .4295
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             <1. .3800
                     /codon_start=3
                     /product="tankyrase-related protein"
                     /protein_id="AAG25674.1"
                     /db_xref="GI:10953952"
                     /translation="HEAGRSWQEGPCQLPPPRRFRTRTADSRCLRRRGAAGGQGAQRG
                     ARVGGHGTAPDPVTAGSQAARALSASSPGGLALLLAGPGLLLRLLALLLAVAAARIMS
                     GRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLVTPEKVNSRDTAGRKST
                     PLHFAAXFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFGHAEVVNLLLRHGADPN
                     ARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTALDLADPSAKAVLTGEY
                     KKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLAAGYNRVKIVQLLLQHG
                     ADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLWQFTPLHEAASKNRVEV
                     CSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFQRPLVAAXAREADVTRIKKH
                     LSLEMVIQASQTHETALHCAAASPYPKRKQICELLLRKGANINEKTKEFLTPLHVASE
                     KAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGF
                     TALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGR
                     QSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGA
                     VVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQD
                     LLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLL
                     QHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKWAPTPLHEAAQKGR
                     TQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGV
                     RSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPG
                     VDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVE
                     RLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIF
                     NRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDER
                     HAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGK
                     SFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGM
                     VDG"
ORIGIN

  Query Match           87.9%;  Score 3338.4;  DB 9;  Length 4295;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 3384;  Conservative    1;  Mismatches    3;  Indels    5;  Gaps    4;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              || |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        500 AGNTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 559

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 619

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 679

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 739

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 799

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 859

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 919

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 979

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        980 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1039

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1040 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1099

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1100 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1159

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1160 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1219

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1220 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1279

Qy       1183 AGCATATGAATTT-AAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1241
              ||||||||||||| |||||||||||||||||||| :|||||||||||||||||||||||||
Db       1280 AGCATATGAATTTCAAAGGCCACTCGTTGCTGCAA-STGCACGAGAAGCTGATGTTACTC 1338

Qy       1242 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1301
              |||||||||||||||||||||||||||||||||||  |||||||| ||||||||||||||
Db       1339 GAATCAAAAAACATCTCTCTCTGGAAATGGTGA--TTCAAGCAT-CTCAAACACATGAAA 1395

Qy       1302 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1396 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1455

Qy       1362 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1456 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1515

Qy       1422 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1516 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1575

Qy       1482 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1576 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1635

Qy       1542 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1636 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1695

Qy       1602 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1696 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1755

Qy       1662 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1721
```

```
Db       1756 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1815
Qy       1722 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1781
Db       1816 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1875
Qy       1782 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1841
Db       1876 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1935
Qy       1842 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1901
Db       1936 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1995
Qy       1902 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1961
Db       1996 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 2055
Qy       1962 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2021
Db       2056 ATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATG 2115
Qy       2022 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAA 2081
Db       2116 AAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAA 2175
Qy       2082 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2141
Db       2176 ATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAG 2235
Qy       2142 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2201
Db       2236 ATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTT 2295
Qy       2202 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2261
Db       2296 CTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAG 2355
Qy       2262 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2321
Db       2356 CAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGA 2415
Qy       2322 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2381
Db       2416 ATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAG 2475
Qy       2382 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2441
Db       2476 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2535
Qy       2442 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2501
Db       2536 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2595
Qy       2502 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2561
Db       2596 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2655
Qy       2562 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2621
Db       2656 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2715
Qy       2622 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2681
Db       2716 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2775
Qy       2682 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2741
Db       2776 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2835
Qy       2742 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2801
Db       2836 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2895
Qy       2802 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2861
Db       2896 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2955
Qy       2862 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2921
Db       2956 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 3015
Qy       2922 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 2981
Db       3016 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 3075
Qy       2982 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3041
Db       3076 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTT 3135
Qy       3042 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3101
Db       3136 TGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTC 3195
Qy       3102 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3161
Db       3196 AGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAG 3255
Qy       3162 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3221
Db       3256 GTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTAT 3315
Qy       3222 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3281
Db       3316 GGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATG 3375
Qy       3282 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3341
Db       3376 AACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATG 3435
Qy       3342 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3401
Db       3436 AAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTT 3495
Qy       3402 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3461
Db       3496 CCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAG 3555
Qy       3462 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3521
Db       3556 ACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGT 3615
Qy       3522 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3581
Db       3616 CTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCA 3675
Qy       3582 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3641
Db       3676 CTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAAC 3735
Qy       3642 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3701
Db       3736 AGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATG 3795
Qy       3702 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3761
Db       3796 GATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTG 3855
Qy       3762 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       3856 GCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3888

RESULT 18
AF342982
```

```
LOCUS       AF342982               3501 bp    mRNA    linear   PRI 18-SEP-2001
DEFINITION  Homo sapiens tankyrase 2 mRNA, complete cds.
ACCESSION   AF342982
VERSION     AF342982.1  GI:13430364
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3501)
  AUTHORS   Kaminker,P.G., Kim,S.H., Taylor,R.D., Zebarjadian,Y., Funk,W.D.,
            Morin,G.B., Yaswen,P. and Campisi,J.
  TITLE     TANK2, a new TRF1-associated poly(ADP-ribose) polymerase, causes
            rapid induction of cell death upon overexpression
  JOURNAL   J. Biol. Chem. 276 (38), 35891-35899 (2001)
  MEDLINE   21443728
   PUBMED   11454873
REFERENCE   2  (bases 1 to 3501)
  AUTHORS   Kaminker,P.G., Campisi,J., Kim,S.H., Yaswen,P. and Morin,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2001) Life Sciences, Lawrence Berkeley National
            Labs, 1 Cyclotron Rd 84-171, Berkeley, CA 94720, USA
FEATURES             Location/Qualifiers
     source          1. .3501
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q23-q24"
     CDS             1. .3501
                     /note="TNKS2; telomere-associated poly(ADP-ribose)
                     polymerase; contains ankyrin domain, sterile alpha motif
                     and PARP catalytic domain; similar to TNKS"
                     /codon_start=1
                     /product="tankyrase 2"
                     /protein_id="AAK25811.1"
                     /db_xref="GI:13430365"
                     /translation="MSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
                     INEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match             86.9%;  Score 3300.4;  DB 9;  Length 3501;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3301;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 557

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 617

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 677

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 737

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 797

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 857

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 917

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 977

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1037

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1038 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1097

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1098 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1157

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1158 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1217

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1218 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1277

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1278 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1337

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1338 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1397

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1398 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1457
Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db       1458 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1517
Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db       1518 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1577
Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db       1578 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1637
Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db       1638 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1697
Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       1698 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1757
Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       1758 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1817
Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db       1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1877
Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       1878 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1937
Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       1938 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 1997
Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       1998 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2057
Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2058 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2117
Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2118 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2177
Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2178 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2237
Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2238 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2297
Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357
Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2358 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2418 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2477
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2478 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2537

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2538 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2597
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2657
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2658 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2717
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2718 GATGGGGCACAAGGAGCTGAAGGAGATTGGWATCAATGCTTATGGACATAGGCACAAACT 2777
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2778 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2837
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       2838 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2897
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       2898 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2957
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       2958 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3017
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3018 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3077
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3078 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3137
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3138 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3197
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3198 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3257
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3258 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3317
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3318 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3377
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3378 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3437
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3438 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3497
Qy       3703 ATAA 3706
Db       3498 ATAA 3501
```

RESULT 19
AX029397
LOCUS AX029397 3400 bp DNA linear PAT 16-SEP-2000

```
DEFINITION  Sequence 1 from Patent WO9915647.
ACCESSION   AX029397
VERSION     AX029397.1  GI:10190194
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Sutherland,R.L. and Daly,R.J.
  TITLE     A potential effector for the grb7 family of signalling proteins
  JOURNAL   Patent: WO 9915647-A 1 01-APR-1999;
            GARVAN INST MED RES (AU) ; SUTHERLAND ROBERT LYNDSAY (AU) ; DALY
            ROGER JOHN (AU)
FEATURES             Location/Qualifiers
     source          1. .3400
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             86.8%;  Score 3297;  DB 6;  Length 3400;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        482 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60

Qy        542 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120

Qy        602 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180

Qy        662 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 721
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240

Qy        722 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300

Qy        782 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360

Qy        842 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 901
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420

Qy        902 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 961
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480

Qy        962 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1021
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540

Qy       1022 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600

Qy       1082 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1141
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660

Qy       1142 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720

Qy       1202 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780

Qy       1262 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840

Qy       1322 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900

Qy       1382 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960

Qy       1442 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020

Qy       1502 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080

Qy       1562 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy       1622 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200

Qy       1682 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260

Qy       1742 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320

Qy       1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380

Qy       1862 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1921
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440

Qy       1922 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500

Qy       1982 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560

Qy       2042 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2101
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620

Qy       2102 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680

Qy       2162 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740

Qy       2222 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800

Qy       2282 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2341
```

```
Db       1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860

Qy       2342 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2401

Db       1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920

Qy       2402 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2461

Db       1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980

Qy       2462 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2521

Db       1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040

Qy       2522 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2581

Db       2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100

Qy       2582 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2641

Db       2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160

Qy       2642 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2701

Db       2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220

Qy       2702 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2761

Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280

Qy       2762 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2821

Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340

Qy       2822 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2881

Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400

Qy       2882 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2941

Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460

Qy       2942 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3001

Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520

Qy       3002 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3061

Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580

Qy       3062 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3121

Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640

Qy       3122 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3181

Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700

Qy       3182 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3241

Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3242 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3301

Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820

Qy       3302 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3361

Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880

Qy       3362 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3421

Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940

Qy       3422 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3481

Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000

Qy       3482 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3541

Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060

Qy       3542 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3601

Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120

Qy       3602 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3661

Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180

Qy       3662 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3721

Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240

Qy       3722 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3781

Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300

Qy       3782 TGCTGAAAAAAAAA 3794

Db       3301 TGCTGAAAAAAAA 3313
```

```
RESULT 20
BD075726
LOCUS       BD075726               3400 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Potential effector of signal transducing protein GRB7 family.
ACCESSION   BD075726
VERSION     BD075726.1  GI:22621329
KEYWORDS    JP 2001517435-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3400)
  AUTHORS   Dary,R.J. and Sutherland,R.L.
  TITLE     Potential effector of signal transducing protein GRB7 family
  JOURNAL   Patent: JP 2001517435-A 1 09-OCT-2001;
            GARVAN INSTITUTE OF MEDICAL RESEARCH
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001517435-A/1
            PD   09-OCT-2001
            PF   23-SEP-1998 JP 2000512939
            PR   23-SEP-1997 AU      PO 9388
            PI   ROGER JOHN DARY,ROBERT LINDSEY SUTHERLAND
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12Q1/68,C12N15/00,C12N5/00
            CC   Potential effector of signal transducing protein GRB7 family
            FH   Key             Location/Qualifiers
            FT   source          1. .3400
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .3400
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           86.8%;  Score 3297;  DB 6;  Length 3400;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3303;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy        482 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 541
```

```
Db        1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60
Qy      542 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 601
Db       61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120
Qy      602 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 661
Db      121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180
Qy      662 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 721
Db      181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240
Qy      722 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 781
Db      241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300
Qy      782 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 841
Db      301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360
Qy      842 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 901
Db      361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420
Qy      902 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 961
Db      421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480
Qy      962 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1021
Db      481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540
Qy     1022 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1081
Db      541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600
Qy     1082 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1141
Db      601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660
Qy     1142 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1201
Db      661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720
Qy     1202 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1261
Db      721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780
Qy     1262 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1321
Db      781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840
Qy     1322 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1381
Db      841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900
Qy     1382 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1441
Db      901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960
Qy     1442 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1501
Db      961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020
Qy     1502 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1561
Db     1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080
Qy     1562 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621
Db     1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy     1622 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1681
Db     1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200
Qy     1682 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1741
Db     1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260
Qy     1742 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1801
Db     1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320
Qy     1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1861
Db     1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380
Qy     1862 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1921
Db     1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440
Qy     1922 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1981
Db     1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500
Qy     1982 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2041
Db     1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560
Qy     2042 CAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2101
Db     1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620
Qy     2102 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2161
Db     1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680
Qy     2162 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2221
Db     1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740
Qy     2222 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2281
Db     1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800
Qy     2282 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2341
Db     1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860
Qy     2342 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2401
Db     1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920
Qy     2402 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2461
Db     1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980
Qy     2462 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2521
Db     1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040
Qy     2522 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2581
Db     2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100
Qy     2582 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2641
Db     2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160
Qy     2642 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2701
Db     2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220
```

```
Qy       2702 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2761
Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280

Qy       2762 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2821
Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340

Qy       2822 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2881
Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400

Qy       2882 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2941
Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460

Qy       2942 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3001
Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520

Qy       3002 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3061
Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580

Qy       3062 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3121
Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640

Qy       3122 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3181
Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700

Qy       3182 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3241
Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3242 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3301
Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820

Qy       3302 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3361
Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880

Qy       3362 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3421
Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940

Qy       3422 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3481
Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000

Qy       3482 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3541
Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060

Qy       3542 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3601
Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120

Qy       3602 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3661
Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180

Qy       3662 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3721
Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240

Qy       3722 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3781
Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300

Qy       3782 TGCTGAAAAAAAA 3794
Db       3301 TGCTGAAAAAAAA 3313
```

```
RESULT 21
AX062142
LOCUS       AX062142               3508 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 1 from Patent WO0100849.
ACCESSION   AX062142
VERSION     AX062142.1  GI:12540045
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 1 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3508
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .>3508
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25052.1"
                     /db_xref="GI:12540046"
                     /db_xref="REMTREMBL:CAC25052"
                     /translation="ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVK
                     RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNAC
                     SFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG
                     RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPL
                     HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM
                     DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFK
                     GHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK
                     GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ
                     TCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV
                     ETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPL
                     HNACSYGHYEVAELLVKHGAVVNVADLWKPTPLHEAAAKGKYEICKLLLQHGADPTKK
                     NRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRH
                     STPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNAC
                     VNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLT
                     AAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSS
                     VVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKE
                     LKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVE
                     EEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANER
                     MLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHK
                     DRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYR
                     GEQAYPEYLITYQIMRPEGMVDG"
ORIGIN

  Query Match           86.8%;  Score 3296.2;  DB 6;  Length 3508;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3298;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447
```

```
Qy       643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db       448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy       703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db       508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy       763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db       568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy       823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db       628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy       883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db       688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy       943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db       748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy      1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db       808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy      1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db       868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

Qy      1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db       928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

Qy      1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db       988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

Qy      1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db      1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

Qy      1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db      1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

Qy      1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db      1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227

Qy      1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db      1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287

Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db      1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347

Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db      1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407

Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db      1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467

Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db      1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527

Qy      1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db      1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1587

Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db      1588 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1647

Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db      1648 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1707

Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db      1708 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1767

Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db      1768 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1827

Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db      1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1887

Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db      1888 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1947

Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db      1948 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2007

Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db      2008 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2067

Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db      2068 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2127

Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db      2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187

Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db      2188 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2247

Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db      2248 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2307

Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db      2308 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367

Qy      2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db      2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2427

Qy      2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db      2428 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2487

Qy      2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db      2488 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2547

Qy      2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db      2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2607

Qy      2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
```

```
Db       2608 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2728 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2787

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2847

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       2848 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2907

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       2908 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2967

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       2968 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3027

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3028 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3087

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3088 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3207

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3208 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3267

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3268 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3327

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3328 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3387

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3388 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3447

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3448 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3507

Qy       3703 A 3703
              |
Db       3508 A 3508
```

```
RESULT 22
AR232162
LOCUS       AR232162                3498 bp    DNA     linear   PAT 20-DEC-2002
DEFINITION  Sequence 4 from patent US 6455290.
ACCESSION   AR232162
VERSION     AR232162.1  GI:27274049
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3498)
  AUTHORS   Berthelsen,J., Toma,S. and Isacchi,A.
  TITLE     Tankyrase homolog protein (THP), nucleic acids, and methods related
            to the same
  JOURNAL   Patent: US 6455290-A 4 24-SEP-2002;
FEATURES             Location/Qualifiers
     source          1. .3498
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             86.7%;  Score 3291.8;  DB 6;  Length 3498;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3290;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 557

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 617

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 677

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 737

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 797

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 857

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 917

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 977

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1037

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
```

```
Db      1038 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1097
Qy      1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db      1098 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1157
Qy      1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db      1158 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1217
Qy      1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db      1218 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1277
Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db      1278 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1337
Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db      1338 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1397
Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db      1398 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1457
Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db      1458 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1517
Qy      1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db      1518 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1577
Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db      1578 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1637
Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db      1638 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1697
Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db      1698 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1757
Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db      1758 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1817
Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db      1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1877
Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db      1878 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1937
Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db      1938 TGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 1997
Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db      1998 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2057
Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db      2058 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2117
Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db      2118 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2177
Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db      2178 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2237
Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db      2238 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGKTGCTAGC 2297
Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db      2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357
Qy      2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db      2358 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417
Qy      2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db      2418 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2477
Qy      2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db      2478 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2537
Qy      2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db      2538 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2597
Qy      2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db      2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGNATCTTGG 2657
Qy      2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db      2658 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2717
Qy      2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db      2718 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2777
Qy      2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db      2778 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2837
Qy      3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db      2838 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2897
Qy      3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db      2898 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2957
Qy      3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db      2958 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3017
Qy      3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db      3018 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3077
Qy      3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db      3078 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3137
Qy      3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db      3138 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3197
Qy      3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db      3198 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3257
```

```
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3258 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3317

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3318 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3377

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3378 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3437

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3438 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3497

Qy       3703 A 3703
              |
Db       3498 A 3498
```

```
RESULT 23
AX076214
LOCUS       AX076214                3498 bp    DNA     linear   PAT 06-FEB-2001
DEFINITION  Sequence 4 from Patent WO0104326.
ACCESSION   AX076214
VERSION     AX076214.1  GI:12710839
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Berthelsen,J., Toma,S. and Isacchi,A.
  TITLE     Tankyrase homolog protein (thp), nucleic acids, and methods related
            to the same
  JOURNAL   Patent: WO 0104326-A 4 18-JAN-2001;
            PHARMACIA & UPJOHN S.p.A. (IT)
FEATURES             Location/Qualifiers
     source          1. .3498
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .>3498
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC28482.1"
                     /db_xref="GI:12710840"
                     /db_xref="REMTREMBL:CAC28482"
                     /translation="MSGRRCAGGGAACASAAABEAVEPAARELFEACRNGDVERVKRLV
                     TPEKVNSRDTAGRKSTPLHLAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
                     HAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
                     LDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
                     AGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
                     QFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFKGHS
                     LLQAAREADVTRIKKHLSLEMVNFKHPQTHETAXHCAAASPYPKRKQICELLLRKGAX
                     INEKTKEFLTPLHVASEKAHNDXVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCR
                     LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETV
                     KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKXXLVPLHNA
                     CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRD
                     GNTPLDLVKDGDTDIQDLLRGDAALLDAAKKSCLARVKKLSSPDNVNCRDTQGRHSTP
                     LHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNA
                     TDKWAFTPLHEAAQKGRTQLCALXLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAM
                     PPSALPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVS
                     SSGTEGASSLEKKEVPGVDFSITQFVRXLGLEHLMDIFEREQITLDVLVEMGHKELKE
                     IGINAYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEM
                     QSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLF
                     HGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRS
                     CYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQ
                     AYPEYLITYQIMRPEGMVDG"
     unsure          368
                     /note="Xaa is Met or Leu or Val"
     unsure          392
                     /note="Xaa is Asn or Thr"
     unsure          415
                     /note="Xaa is Val or Ile"
     unsure          558
                     /note="Xaa is Gly or Glu"
     unsure          559
                     /note="Xaa is Gly or Ala"
     unsure          764
                     /note="Xaa is Leu or Val"
     unsure          884
                     /note="Xaa is Asn or His or Asp or Tyr"
ORIGIN
```

```
  Query Match          86.7%;  Score 3291.8;  DB 6;  Length 3498;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3290;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 557

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 617

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 677

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 737

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 797

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        798 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 857

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        858 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 917

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        918 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 977

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1037
```

```
Qy    1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1038 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1097

Qy    1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1098 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1157

Qy    1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
           ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db    1158 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1217

Qy    1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
           ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db    1218 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1277

Qy    1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1278 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1337

Qy    1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1338 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1397

Qy    1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1398 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1457

Qy    1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1458 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1517

Qy    1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1518 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1577

Qy    1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1578 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1637

Qy    1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
           ||||||||||||||||||||||||||||||||||||||||||::|:||||||||||||||||
Db    1638 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1697

Qy    1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1698 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1757

Qy    1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1758 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1817

Qy    2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1877

Qy    2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1878 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1937

Qy    2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1938 TGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 1997

Qy    2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1998 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2057

Qy    2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2058 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2117

Qy    2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2118 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2177

Qy    2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2178 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2237

Qy    2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||
Db    2238 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGKTGCTAGC 2297

Qy    2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357

Qy    2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2358 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417

Qy    2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2418 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2477

Qy    2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2478 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2537

Qy    2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2538 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2597

Qy    2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGNATCTTGG 2657

Qy    2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2658 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2717

Qy    2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2718 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2777

Qy    2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2778 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2837

Qy    3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2838 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2897

Qy    3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2898 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2957

Qy    3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2958 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3017

Qy    3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3018 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3077

Qy    3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3078 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3137

Qy    3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3138 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3197

Qy    3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
```

```
Db       3198 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3257

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3258 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3317

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3318 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3377

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3378 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3437

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3438 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3497

Qy       3703 A 3703
              |
Db       3498 A 3498


RESULT 24
AR361484
LOCUS       AR361484                4493 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 1 from patent US 6599728.
ACCESSION   AR361484
VERSION     AR361484.1  GI:33769332
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4493)
  AUTHORS   Morin,G.B., Funk,W.D. and Piatyszek,M.A.
  TITLE     Second mammalian tankyrase
  JOURNAL   Patent: US 6599728-A 1 29-JUL-2003;
FEATURES             Location/Qualifiers
     source          1. .4493
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           78.2%;  Score 2969.6;  DB 6;  Length 4493;
  Best Local Similarity  88.6%;  Pred. No. 0;
  Matches 3028;  Conservative    1;  Mismatches  363;  Indels   27;  Gaps    2;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||||||||||| || |  || ||||||||||||| |||||||||| | ||||| |
Db        672 AGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGCTTCAAAATGGTGCAAATNTCCAANC 731

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              || | || |||||||||||||||||||||||| ||||||||||||||||||||||| ||
Db        732 ACTTTATAATGGGGGCCTTATTCCTCTTCATANTGCATGCTCTTTTGGTCATGCTAAAAN 791

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 TATCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 851

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 911

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 971

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1031

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822

Db       1032 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1091

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1151

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1211

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1271

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1331

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1391

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1451

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1452 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1511

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1512 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1571

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1572 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1631

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1632 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1691

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1692 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1751

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1752 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1811

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1812 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1871

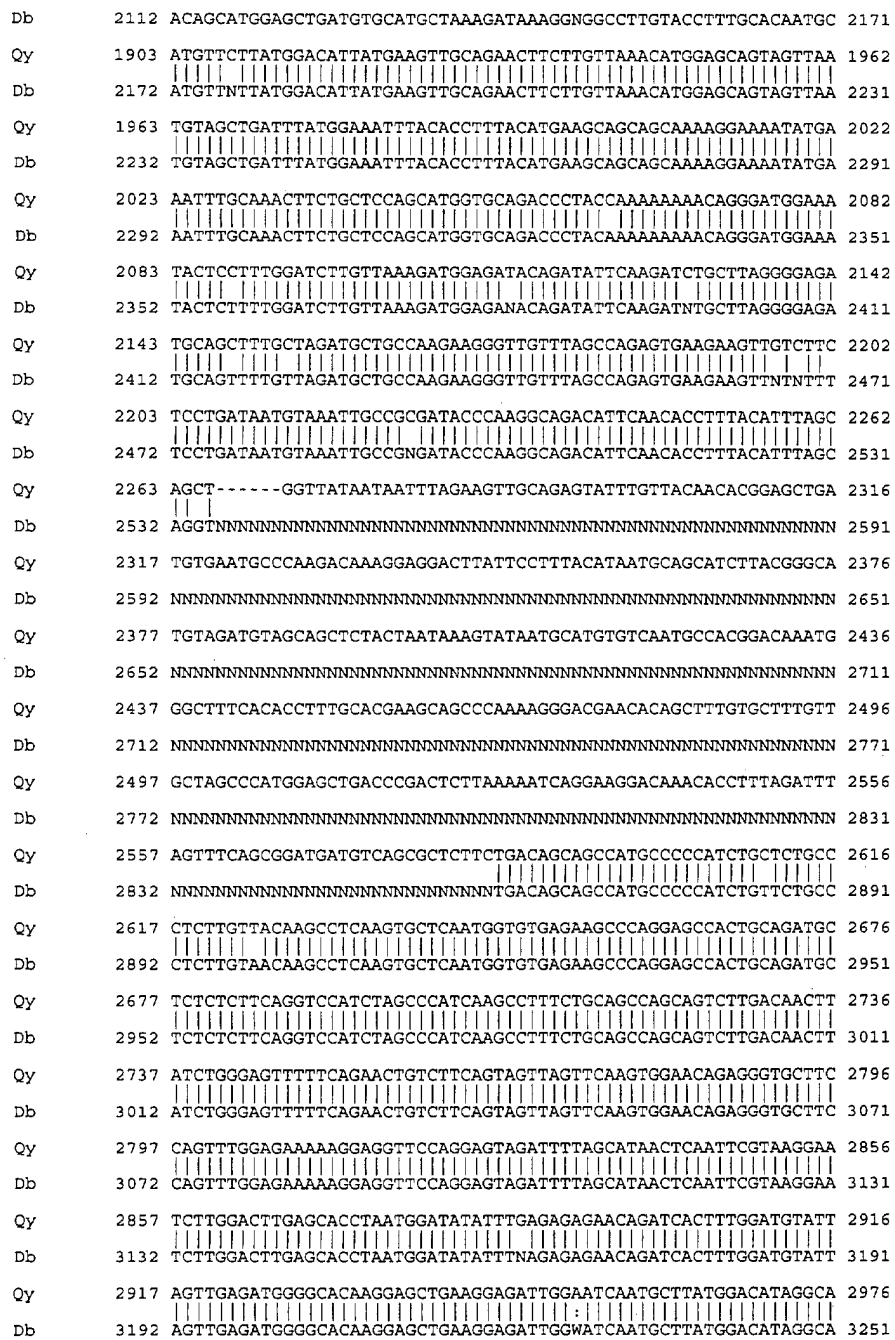
Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1872 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1931

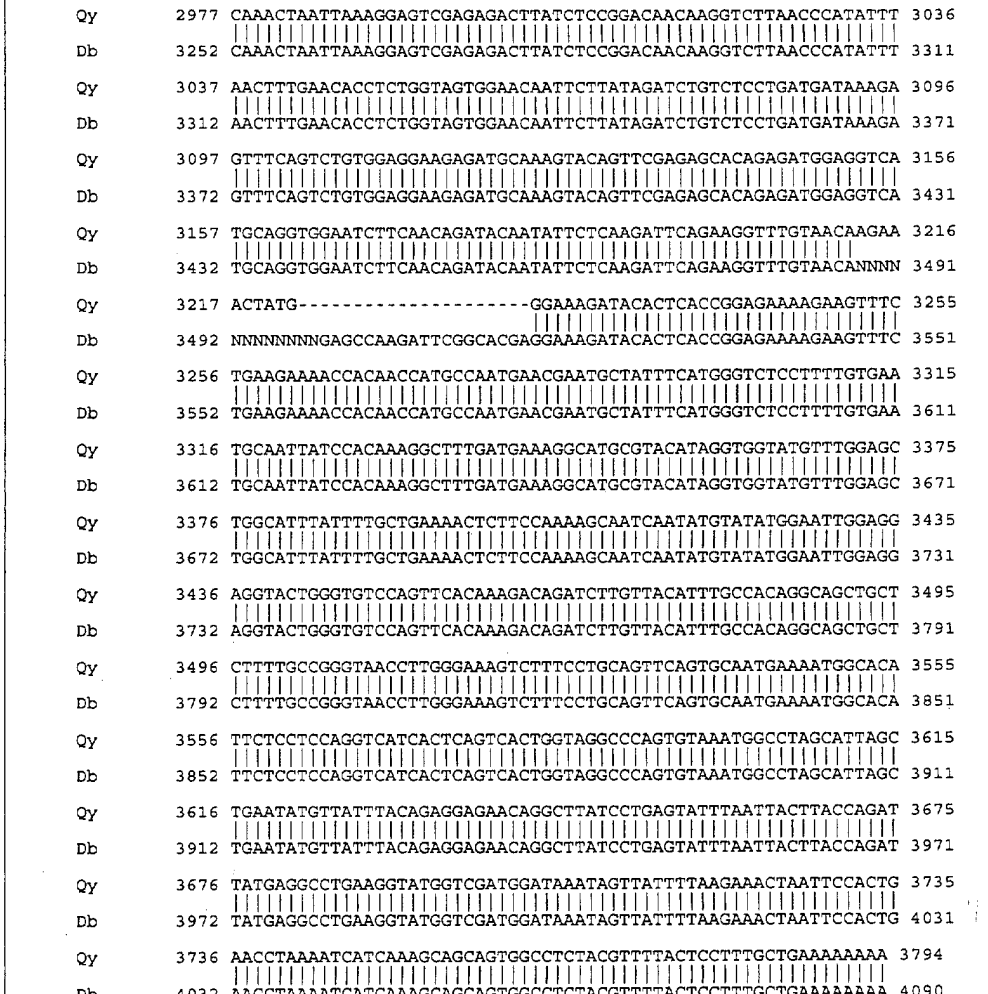
Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1932 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1991

Qy       1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1992 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2051

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2052 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2111

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
```





RESULT 25
AR361485
LOCUS AR361485 4297 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 3 from patent US 6599728.
ACCESSION AR361485
VERSION AR361485.1 GI:33769333
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4297)
AUTHORS Morin,G.B., Funk,W.D. and Piatyszek,M.A.
TITLE Second mammalian tankyrase
JOURNAL Patent: US 6599728-A 3 29-JUL-2003;
FEATURES Location/Qualifiers

```
     source          1. .4297
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             78.2%;  Score 2968.2;  DB 6;  Length 4297;
  Best Local Similarity   88.3%;  Pred. No. 0;
  Matches 3020;  Conservative    7;  Mismatches  381;  Indels   12;  Gaps    2;

Qy        384 CCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGA 443
              |||    ||  | :|    | |||||||||||||||| || ||:|||||||||||| |
Db        476 CCCCGCTGCACTTYCCCGCANGTTTTGGGCGGAAAGACTTANTTRAATATTTGCTTCANA 535

Qy        444 ATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCT 503
              |||||||||  |||:||| ||||||||  ||||||||||||||||||||||||||||||
Db        536 ATGGTGCAAATGTYCAANCACGTGATNATGGGGGCCTTATTCCTCTTCATAATGCATGCT 595

Qy        504 CTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTC 563
              |||||||||:||||:||    || |||||||||||  |||| |||||  |:||||||||
Db        596 CTTTTGGTCMTGCTRAAANTATCNATCTCCTTTTGCNACATNGTGCANAMCCCAATGCTC 655

Qy        564 GAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTT 623
              |||||||||||||||||||||||  | ||||||||||||||||||||||||||    ||
Db        656 GAGATAATTGGAATTATACTCCTCNCNATGAAGCTGCAATTAAAGGAAAGATTGANNNTT 715

Qy        624 GCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAG 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        716 GCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAG 775

Qy        684 CATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATG 743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        776 CATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATG 835

Qy        744 AACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCAT 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        836 AACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCAT 895

Qy        804 TAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAG 863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        896 TAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAG 955

Qy        864 GATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTA 923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        956 GATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTA 1015

Qy        924 AAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAA 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1016 AAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAA 1075

Qy        984 CTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTC 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1076 CTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTC 1135

Qy       1044 CTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATG 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1136 CTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATG 1195

Qy       1104 GTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACAC 1163
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1196 GTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACAC 1255

Qy       1164 CACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAC 1223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1256 CACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCAC 1315

Qy       1224 GAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGC 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1316 GAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGC 1375

Qy       1284 ATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAA 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1376 ATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAA 1435

Qy       1344 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1436 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1495

Qy       1404 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1496 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1555

Qy       1464 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1556 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1615

Qy       1524 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1616 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1675

Qy       1584 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 1643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1676 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 1735

Qy       1644 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1736 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1795

Qy       1704 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1796 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1855

Qy       1764 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1856 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1915

Qy       1824 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1883
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1916 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGNGGCC 1975

Qy       1884 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1943
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1976 TTGTACCTTTGCACAATGCATGTTNTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 2035

Qy       1944 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2036 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2095

Qy       2004 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2063
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       2096 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 2155

Qy       2064 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2123
              ||||||| |   | ||||| | || ||||||||||||||||||||| ||||||||||
Db       2156 AAAAAAAAAAAAAAAGGAAANATTCNTTTGGATCTTGTTAAAGATGGAGANACAGATATTC 2215

Qy       2124 AAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCA 2183
              ||||| |||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db       2216 AAGATNTGCTTAGGGGAGATGCAGTTTTGTTAGATGCTGCCAAGAAGGGTTGTTTAGCCA 2275

Qy       2184 GAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATT 2243
              |||||||||||||  || ||||||||||||||||||| ||||||||||||||||||||||
Db       2276 GAGTGAAGAAGTTNTNTTTTCCTGATAATGTAAATTGCCGNGATACCCAAGGCAGACATT 2335

Qy       2244 CAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTAC 2303
              ||||||||||||||||||  |||
Db       2336 CAACACCTTTACATTTAGCA---GGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2392

Qy       2304 AACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 2363

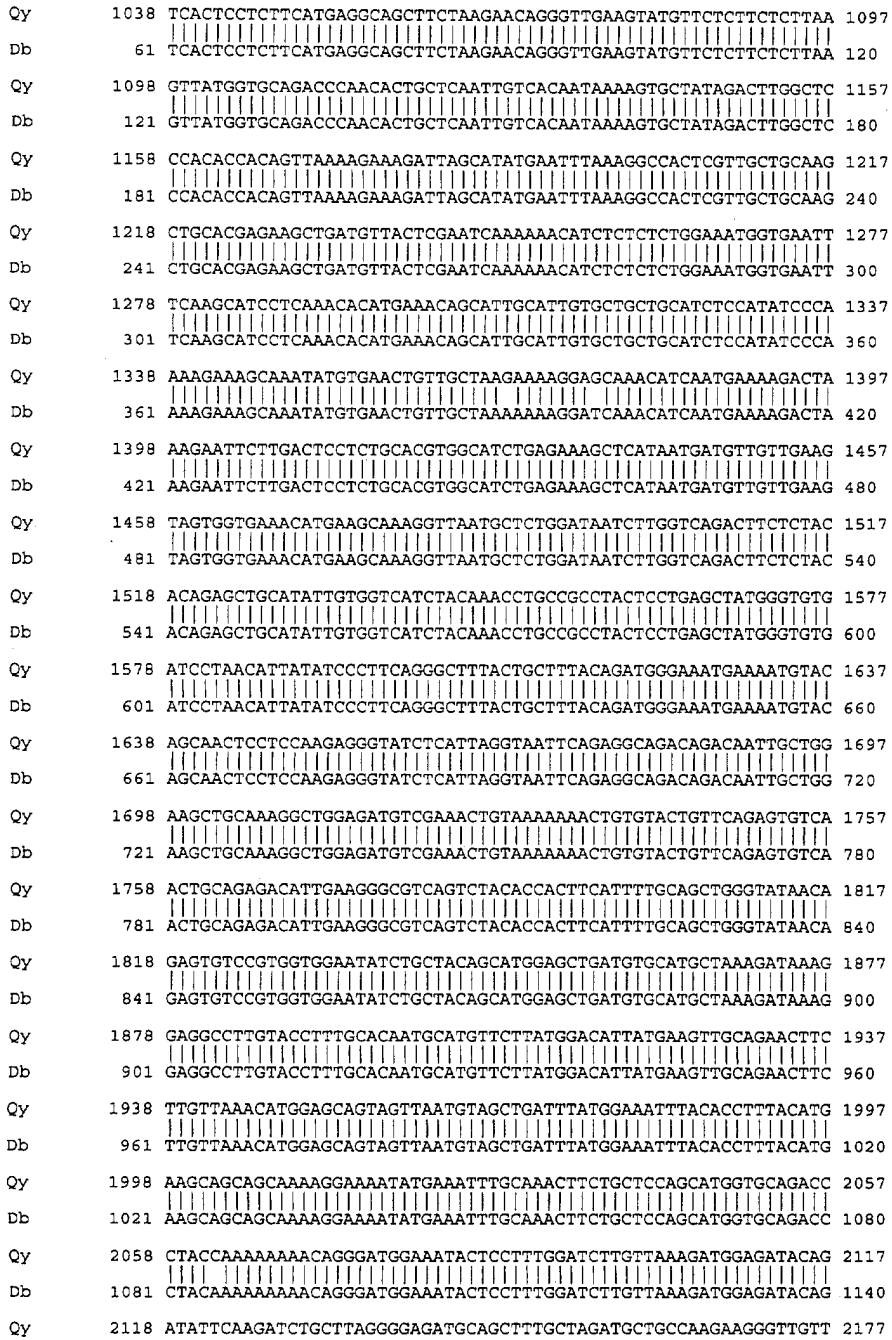
Db       2393 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2452

Qy       2364 CATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATG 2423

Db       2453 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2512

Qy       2424 CCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGC 2483
```

```
Db       2513 NNNN
```


Qy 1038 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 1097
Db 61 TCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAA 120
Qy 1098 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 1157
Db 121 GTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTC 180
Qy 1158 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 1217
Db 181 CCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAG 240
Qy 1218 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 1277
Db 241 CTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATT 300
Qy 1278 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 1337
Db 301 TCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCA 360
Qy 1338 AAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTA 1397
Db 361 AAAGAAAGCAAATATGTGAACTGTTGCTAAAAAAAGGATCAAACATCAATGAAAAGACTA 420
Qy 1398 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 1457
Db 421 AAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAG 480
Qy 1458 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 1517
Db 481 TAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTAC 540
Qy 1518 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 1577
Db 541 ACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTG 600
Qy 1578 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1637
Db 601 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 660
Qy 1638 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1697
Db 661 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 720
Qy 1698 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1757
Db 721 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 780
Qy 1758 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1817
Db 781 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 840
Qy 1818 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1877
Db 841 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 900
Qy 1878 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1937
Db 901 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 960
Qy 1938 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1997
Db 961 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1020
Qy 1998 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2057
Db 1021 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 1080
Qy 2058 CTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2117
Db 1081 CTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 1140
Qy 2118 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2177
Db 1141 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 1200
Qy 2178 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2237
Db 1201 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 1260
Qy 2238 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 2297
Db 1261 GACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATT 1320
Qy 2298 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 2357
Db 1321 TGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATA 1380
Qy 2358 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 2417
Db 1381 ATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTG 1440
Qy 2418 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 2477
Db 1441 TCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAA 1500
Qy 2478 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 2537
Db 1501 CACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAG 1560
Qy 2538 GACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCA 2597
Db 1561 GACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCA 1620
Qy 2598 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 2657
Db 1621 TGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCC 1680
Qy 2658 CAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 2717
Db 1681 CAGGAGCCACTGCAGATGCTCTCT-TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAG 1739
Qy 2718 CCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 2777
Db 1740 CCAGCAGTCTTGACAAC-TATCTGGGAG-TTTTCAGAACTGTCTTCAGTAGTTAGTTCAA 1797
Qy 2778 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 2837
Db 1798 GTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCA 1857
Qy 2838 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 2897
Db 1858 TAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAAC 1917
Qy 2898 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 2957
Db 1918 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 1977
Qy 2958 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 3017
Db 1978 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 2037
Qy 3018 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 3077
Db 2038 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 2097
Qy 3078 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 3137
Db 2098 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 2157
Qy 3138 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 3197
Db 2158 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 2217
Qy 3198 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 3257

```
Db       2218 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 2277

Qy       3258 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 3317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2278 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 2337

Qy       3318 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2338 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 2397

Qy       3378 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 3437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2398 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 2457

Qy       3438 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 3497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2458 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 2517

Qy       3498 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 3557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2518 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 2577

Qy       3558 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 3617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2578 CTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTG 2637

Qy       3618 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 3677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2638 AATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTA 2697

Qy       3678 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 3737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2698 TGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAA 2757

Qy       3738 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2758 CCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2814


RESULT 27
AX062233
LOCUS       AX062233               2971 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 92 from Patent WO0100849.
ACCESSION   AX062233
VERSION     AX062233.1  GI:12540134
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 92 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .2971
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             64.3%;  Score 2439.8;  DB 6;  Length 2971;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2441;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1352 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1412 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1472 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1532 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1591
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1592 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1652 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1712 GGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1772 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1832 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1892 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1952 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2012 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720

Qy       2072 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2132 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2192 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2252 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2372 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2432 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy       2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200
```

```
Qy       2552 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2611
             ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2612 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy       2672 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy       2732 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy       2792 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy       2852 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy       2912 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2972 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy       3032 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy       3092 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy       3152 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860

Qy       3212 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy       3272 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy       3332 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy       3392 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy       3452 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy       3512 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3572 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3692 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy       3752 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             |||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 28
AX062229
LOCUS AX062229 3353 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 88 from Patent WO0100849.
ACCESSION AX062229
VERSION AX062229.1 GI:12540130
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 88 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .3353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS <1. .2355
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC25057.1"
/db_xref="GI:12540131"
/db_xref="REMTREMBL:CAC25057"
/translation="CELLLRKGANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKV
NALDNLGQTSLHRAAYCGHLQTCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEG
ISLGNSEADRQLLEAAKAGDVETVKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVV
EYLLQHGADVHAKDKGGLVPLHNACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA
AKGKYEICKLLLQHGADPTKKNRDGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCL
ARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPL
HNAASYGHVDVAALLIKYNACVNATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLK
NQEGQTPLDLVSADDVSALLTAAMPPSALPSCYKPQVLNGVRSPGATADALSSGPSSP
SSLSAASSLDNLSGSFSELSSVVSSSGTEGASSLEKKEVPGVDFSITQFVRNLGLEHL
MDIFEREQITLDVLVEMGHKELKEIGINAYGHRHKLIKGVERLISGQQGLNPYLTLNT
SGSGTILIDLSPDDKEFQSVEEEMQSTVREHRDGGHAGGIFNRYNILKIQKVCNKKLW
ERYTHRRKEVSEENHNHANERMLFHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAEN
SSKSNQYVYGIGGGTGCPVHKDRSCYICHRQLLFCRVTLGKSFLQFSAMKMAHSPPGH
HSVTGRPSVNGLALAEYVIYRGEQAYPEYLITYQIMRPEGMVDG"
ORIGIN

Query Match 64.3%; Score 2439.8; DB 6; Length 3353;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1352 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1412 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1472 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1532 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240
Qy       1592 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1651
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300
Qy       1652 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1711
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360
Qy       1712 GGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1771
Db        361 GGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420
Qy       1772 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1831
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480
Qy       1832 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1891
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540
Qy       1892 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1951
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600
Qy       1952 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660
Qy       2012 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2071
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720
Qy       2072 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2131
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780
Qy       2132 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2191
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840
Qy       2192 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2251
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900
Qy       2252 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2311
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960
Qy       2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2371
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020
Qy       2372 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2431
Db       1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080
Qy       2432 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2491
Db       1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140
Qy       2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2551
Db       1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200
Qy       2552 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2611
Db       1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260
Qy       2612 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2671
Db       1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320
```

```
Qy       2672 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2731
Db       1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380
Qy       2732 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2791
Db       1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440
Qy       2792 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2851
Db       1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500
Qy       2852 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2911
Db       1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560
Qy       2912 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2971
Db       1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620
Qy       2972 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3031
Db       1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680
Qy       3032 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3091
Db       1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740
Qy       3092 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3151
Db       1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800
Qy       3152 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3211
Db       1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860
Qy       3212 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3271
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920
Qy       3272 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3331
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980
Qy       3332 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3391
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040
Qy       3392 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100
Qy       3452 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3511
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160
Qy       3512 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3571
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220
Qy       3572 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3631
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280
Qy       3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3691
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340
Qy       3692 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3751
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400
```

```
Qy       3752 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
              ||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443


RESULT 29
AY142107
LOCUS       AY142107                3930 bp    mRNA    linear   VRT 30-MAY-2003
DEFINITION  Gallus gallus tankyrase 2 mRNA, complete cds.
ACCESSION   AY142107
VERSION     AY142107.1  GI:27461952
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 3930)
  AUTHORS   De Rycker,M., Venkatesan,R.N., Wei,C. and Price,C.M.
  TITLE     Vertebrate tankyrase domain structure and sterile alpha motif
            (SAM)-mediated multimerization
  JOURNAL   Biochem. J. 372 (Pt 1), 87-96 (2003)
  MEDLINE   22615216
   PUBMED   12589701
REFERENCE   2  (bases 1 to 3930)
  AUTHORS   De Rycker,M., Venkatesan,R.N. and Price,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2002) Department of Molecular Genetics,
            University of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH
            45267-0524, USA
FEATURES             Location/Qualifiers
     source          1. .3930
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
     CDS             171. .3674
                     /codon_start=1
                     /product="tankyrase 2"
                     /protein_id="AAN41650.1"
                     /db_xref="GI:27461953"
                     /translation="MAARRCAGGAAALAEAPGCGSAVEPARELFEACRNGDVERVKRL
                     VRPENVNSRDTAGRKSSPLHFAAGFGRKDVVEYLLQSGANVHARDDGGLIPLHNACSF
                     GHAEVVNLLLRHGANPNARDNWNYTPLHEAAIKGKTDVCIVLLQHGAEPTIRNTDGRT
                     ALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMSLLTPLNVNCHASDGRKSTPLHL
                     AAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDL
                     WQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSTIDLAPTPQLKERLAYEFKGH
                     SLLQAARESDVARIKKHLSLETVNFKHPQTHETALHCAAASPYPKRKQVCELLLRKGA
                     NINEKTKDFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAHCGHLQTC
                     RLLLSSGCDPSIVSLQGFTALQMGNESVQQLLQEGIPLGNSDADRQLLEAAKAGDVDT
                     VKKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHN
                     ACSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPPKKNR
                     DGNTPLDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLCSPDNVNCRDTQGRHST
                     PLHLAAGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVN
                     ATDKWAFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVTADDVSALLTAA
                     MPPSALPSCYKPQVISVSQTASSTADPLSSVPSSPSSLSAASSLDNLSGSFSELPSVV
                     GTNSAEGATVLEKKEVSGVDFSINQFVRNLGLEHLIDIFEREQITLDVLVEMGHKELK
                     EIGINAYGHRHKIIKGVERLISGQQGLNPYLTLNTSSSGTLLIDLSSEDKEFQSVEEE
                     MQSTVREHRDGGHAGGVFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERML
                     FHGSPFVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPIHKDR
                     SCYVCHRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGE
                     QAYPEYLITYQIVKPEATTEA"
ORIGIN

  Query Match           63.3%;  Score 2403.2;  DB 5;  Length 3930;
  Best Local Similarity  83.0%;  Pred. No. 0;
  Matches 2741;  Conservative    0;  Mismatches  563;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||| ||||| || ||||| ||||| || || |||| |||||| | |||||| ||
Db        371 AGGTTTTGGTCGGAAGGATGTAGTCGAATACTTACTGCAGAGTGGTGCCAATGTCCACGC 430

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||| |||||||| |||||||| || ||||| ||||| ||||||||||||||| ||||||||

Db        431 ACGAGATGATGGAGGCCTTATCCCACTTCACAATGCTTGCTCTTTTGGTCACGCTGAAGT 490

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              |||||||||||  |||| || |||||| | || |||||||| || |||||||||| ||
Db        491 TGTCAATCTCCTCCTGCGGCACGGTGCAAATCCTAATGCTCGGGACAATTGGAATTACAC 550

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||| || ||||| || || ||||| || || | ||||| ||||||||  || | |||||
Db        551 TCCCCTTCATGAGGCAGCCATTAAGGGCAAAACAGATGTCTGCATTGTATTGCTGCAGCA 610

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||| |||||  |||||||||| || |||||||||| ||||| ||||||||||||||| ||
Db        611 TGGTGCTGAACCAACCATCCGGAACACAGATGGAAGAACAGCTTTGGATTTAGCAGACCC 670

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||  |||||||||| |||||||||| |||||||||| |||||||||||||||||
Db        671 ATCTGCAAAAGCAGTGCTGACTGGTGAATACAAGAAAGATGAGCTCTTAGAAAGTGCCAG 730

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              |||||| |||||||||| |||||| |||| || |||||||| ||||| ||||| || ||
Db        731 GAGTGGAAATGAAGAAAAGATGATGTCTCTTCTTACACCATTGAATGTTAACTGTCATGC 790

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              |||||||||| ||||| ||||  ||||| | ||||| || |||||| | ||||| ||
Db        791 GAGTGATGGCAGGAAGTCTACTCCTCTACATCTAGCAGCTGGGTATAACCGTGTAAAAAT 850

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              |||||||||||||||||||||||||||||||||||| |||| ||||| || ||| ||||
Db        851 AGTACAGCTGTTACTGCAACATGGGGCTGATGTGCACGCTAAGGATAAGGGAGATTTGGT 910

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||| | |||||||||||||| ||||| |||||||||||||| || ||| | || |||||
Db        911 GCCACTTCACAATGCCTGTTCCTATGGCCATTATGAAGTAACAGAGCTTCTAGTGAAGCA 970

Qy        1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
               ||| || ||||| ||||| |||||  ||||||| |||||||||| ||||| ||||||||
Db        971 TGGAGCATGTGTGAATGCGATGGATCTGTGGCAGTTCACTCCTCTGCATGAAGCAGCTTC 1030

Qy        1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
               |||||||||| |||||||| |||||| | |||||||||||||| ||||||||||| |
Db        1031 AAAGAACAGGGTGGAAGTATGCTCTCTTTTGTTAAGTTATGGTGCTGACCCAACACTGTT 1090

Qy        1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
               || |||||||| |||||  | || || |||||||| || || || ||||||||| ||||
Db        1091 GAACTGTCACAACAAAAGCACCATTGATTTGGCTCCAACGCCTCAATTAAAAGAACGATT 1150

Qy        1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
               |||||||||| |||||  |||| |||||||||| |||||  |||| | |||||| ||||
Db        1151 AGCATATGAATTCAAAGGTCACTCACTGCTACAAGCTGCAAGAGAATCAGATGTTGCTCG 1210

Qy        1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
               || || || ||||| ||| |||||  || || |||||||||||||||| ||||| ||
Db        1211 TATAAAGAAGCATCTTTCTTTGGAAACAGTTAACTTCAAGCATCCTCAAACGCATGAGAC 1270

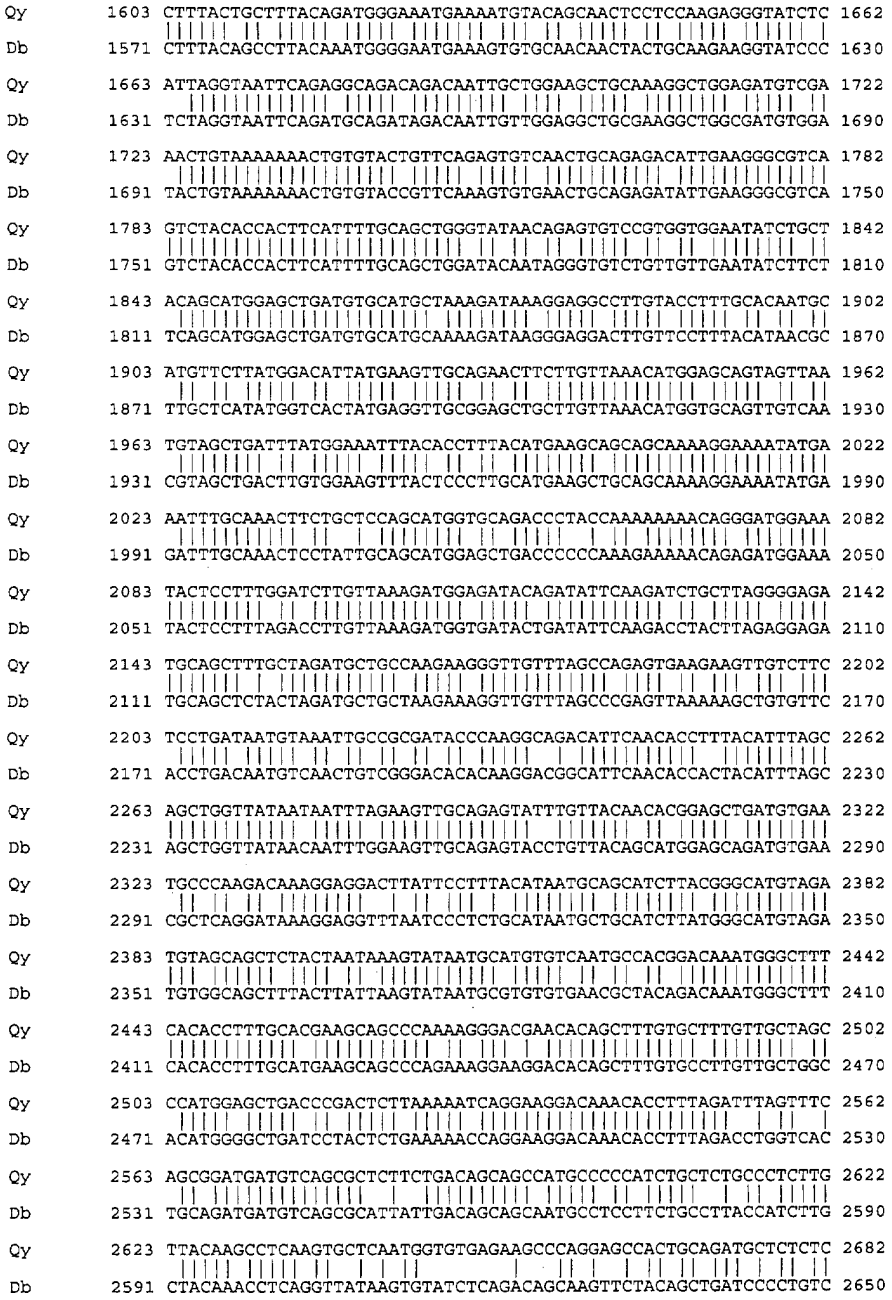
Qy        1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
               |||| |||||||||||||||||||||| ||||| || ||||| ||| |||||||| || |
Db        1271 AGCACTGCACTGTGCTGCTGCATCTCCTTATCCTAAGAGAAAACAAGTATGTGAATTGCT 1330

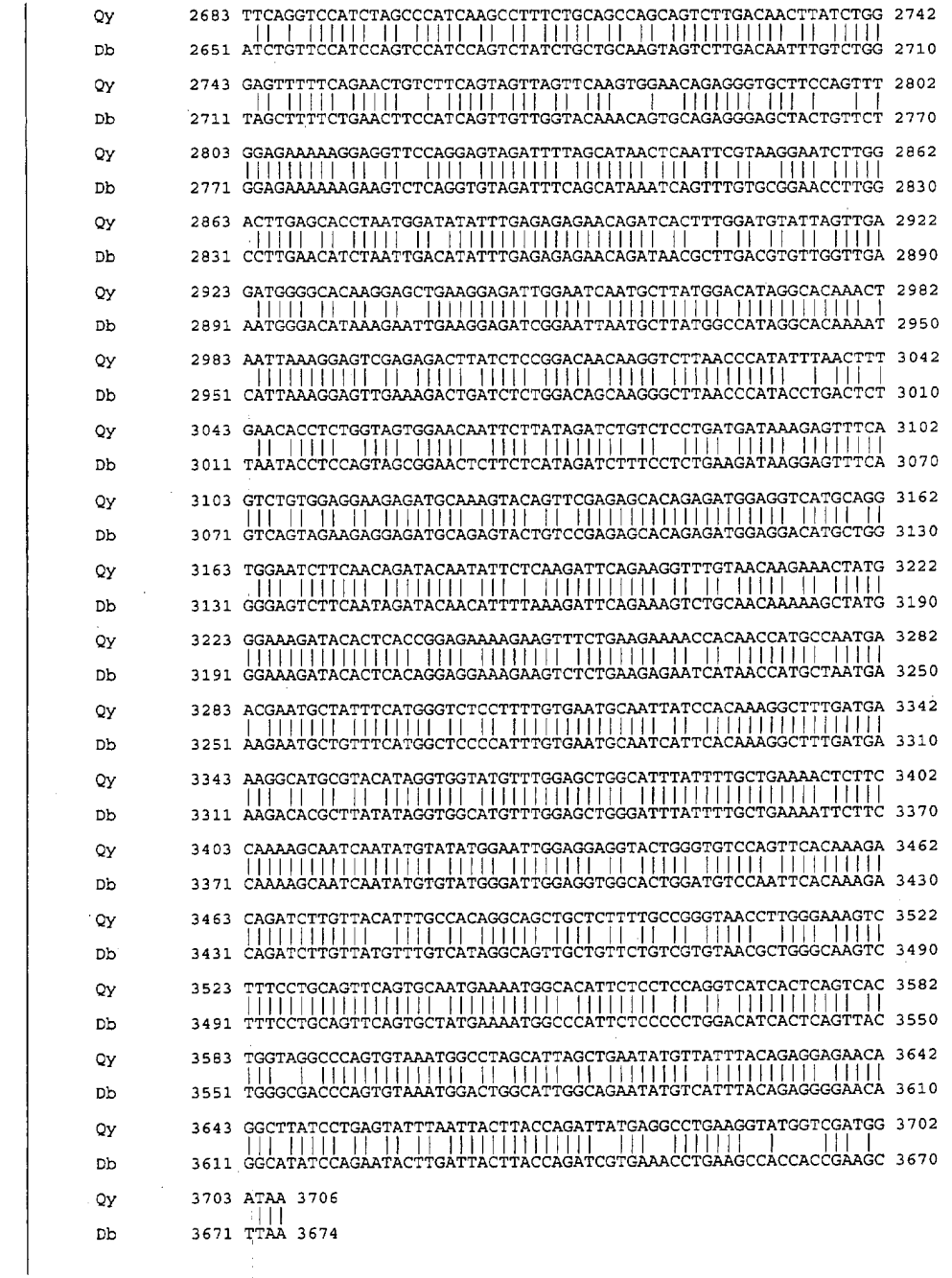
Qy        1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
               ||| || |||||||| ||||| |||||||||| ||||| ||||||||||||||||| ||
Db        1331 GCTGAGGAAAGGAGCTAACATTAATGAAAAGACAAAAGACTTCTTGACTCCTCTGCATGT 1390

Qy        1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
               |||||| || |||||||||||||||||||||||||||||| ||||||||||||||||||
Db        1391 GGCATCAGAAAAAGCTCATAATGATGTTGTTGAAGTAGTTGTGAAACATGAAGCAAAGGT 1450

Qy        1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
               |||||| |||||| |||| |||||| ||||| ||||||||| ||| |||||||||| ||
Db        1451 TAATGCATTGGATAACCTTGGCCAGACCTCTCTTCACAGAGCGGCACATTGTGGTCACCT 1510

Qy        1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
               || || ||| |  |  | |||| | ||| |||||||||  |||| | ||  | |||||
Db        1511 CCAGACTTGCAGGTTGTTGTTGAGTTCTGGATGTGATCCTTCCATTGTGTCTTTGCAGGG 1570
```


Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1571 CTTTACAGCCTTACAAATGGGGAATGAAAGTGTGCAACAACTACTGCAAGAAGGTATCCC 1630
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1631 TCTAGGTAATTCAGATGCAGATAGACAATTGTTGGAGGCTGCGAAGGCTGGCGATGTGGA 1690
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1691 TACTGTAAAAAAACTGTGTACCGTTCAAAGTGTGAACTGCAGAGATATTGAAGGGCGTCA 1750
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1751 GTCTACACCACTTCATTTTGCAGCTGGATACAATAGGGTGTCTGTTGTTGAATATCTTCT 1810
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1811 TCAGCATGGAGCTGATGTGCATGCAAAAGATAAGGGAGGACTTGTTCCTTTACATAACGC 1870
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1871 TTGCTCATATGGTCACTATGAGGTTGCGGAGCTGCTTGTTAAACATGGTGCAGTTGTCAA 1930
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1931 CGTAGCTGACTTGTGGAAGTTTACTCCCTTGCATGAAGCTGCAGCAAAAGGAAAATATGA 1990
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 1991 GATTTGCAAACTCCTATTGCAGCATGGAGCTGACCCCCCAAAGAAAAACAGAGATGGAAA 2050
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2051 TACTCCTTTAGACCTTGTTAAAGATGGTGATACTGATATTCAAGACCTACTTAGAGGAGA 2110
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2111 TGCAGCTCTACTAGATGCTGCTAAGAAAGGTTGTTTAGCCCGAGTTAAAAAGCTGTGTTC 2170
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2171 ACCTGACAATGTCAACTGTCGGGACACACAAGGACGGCATTCAACACCACTACATTTAGC 2230
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2231 AGCTGGTTATAACAATTTGGAAGTTGCAGAGTACCTGTTACAGCATGGAGCAGATGTGAA 2290
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2291 CGCTCAGGATAAAGGAGGTTTAATCCCTCTGCATAATGCTGCATCTTATGGGCATGTAGA 2350
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2351 TGTGGCAGCTTTACTTATTAAGTATAATGCGTGTGTGAACGCTACAGACAAATGGGCTTT 2410
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2411 CACACCTTTGCATGAAGCAGCCCAGAAAGGAAGGACACAGCTTTGTGCCTTGTTGCTGGC 2470
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2471 ACATGGGGCTGATCCTACTCTGAAAAACCAGGAAGGACAAACACCTTTAGACCTGGTCAC 2530
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2531 TGCAGATGATGTCAGCGCATTATTGACAGCAGCAATGCCTCCTTCTGCCTTACCATCTTG 2590
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2591 CTACAAACCTCAGGTTATAAGTGTATCTCAGACAGCAAGTTCTACAGCTGATCCCCTGTC 2650


Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2651 ATCTGTTCCATCCAGTCCATCCAGTCTATCTGCTGCAAGTAGTCTTGACAATTTGTCTGG 2710
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2711 TAGCTTTTCTGAACTTCCATCAGTTGTTGGTACAAACAGTGCAGAGGGAGCTACTGTTCT 2770
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2771 GGAGAAAAAAGAAGTCTCAGGTGTAGATTTCAGCATAAATCAGTTTGTGCGGAACCTTGG 2830
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2831 CCTTGAACATCTAATTGACATATTTGAGAGAGAACAGATAACGCTTGACGTGTTGGTTGA 2890
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2891 AATGGGACATAAAGAATTGAAGGAGATCGGAATTAATGCTTATGGCCATAGGCACAAAAT 2950
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2951 CATTAAAGGAGTTGAAAGACTGATCTCTGGACAGCAAGGGCTTAACCCATACCTGACTCT 3010
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3011 TAATACCTCCAGTAGCGGAACTCTTCTCATAGATCTTTCCTCTGAAGATAAGGAGTTTCA 3070
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3071 GTCAGTAGAAGAGGAGATGCAGAGTACTGTCCGAGAGCACAGAGATGGAGGACATGCTGG 3130
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3131 GGGAGTCTTCAATAGATACAACATTTTAAAGATTCAGAAAGTCTGCAACAAAAAGCTATG 3190
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3191 GGAAAGATACACTCACAGGAGGAAAGAAGTCTCTGAAGAGAATCATAACCATGCTAATGA 3250
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3251 AAGAATGCTGTTTCATGGCTCCCCATTTGTGAATGCAATCATTCACAAAGGCTTTGATGA 3310
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3311 AAGACACGCTTATATAGGTGGCATGTTTGGAGCTGGGATTTATTTTGCTGAAAATTCTTC 3370
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3371 CAAAAGCAATCAATATGTGTATGGGATTGGAGGTGGCACTGGATGTCCAATTCACAAAGA 3430
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3431 CAGATCTTGTTATGTTTGTCATAGGCAGTTGCTGTTCTGTCGTGTAACGCTGGGCAAGTC 3490
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3491 TTTCCTGCAGTTCAGTGCTATGAAAATGGCCCATTCTCCCCCTGGACATCACTCAGTTAC 3550
Qy 3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db 3551 TGGGCGACCCAGTGTAAATGGACTGGCATTGGCAGAATATGTCATTTACAGAGGGGAACA 3610
Qy 3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db 3611 GGCATATCCAGAATACTTGATTACTTACCAGATCGTGAAACCTGAAGCCACCACCGAAGC 3670
Qy 3703 ATAA 3706
Db 3671 TTAA 3674

```
RESULT 30
AX062231
LOCUS       AX062231               3799 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 90 from Patent WO0100849.
ACCESSION   AX062231
VERSION     AX062231.1  GI:12540132
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 90 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3799
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <3. .2273
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25058.1"
                     /db_xref="GI:12540133"
                     /db_xref="REMTREMBL:CAC25058"
                     /translation="AHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQTCRLLLS
                     YGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDVETVKKLC
                     TVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNACSYG
                     HYEVAELLVKHGAVVNVADLWKFTPLHEAAAKGKYEICKLLLQHGADPTKKNRDGNTP
                     LDLVKDGDTDIQDLLRGDAALLDAAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLA
                     AGYNNLEVAEYLLQHGADVNAQDKGGLIPLHNAASYGHVDVAALLIKYNACVNATDKW
                     AFTPLHEAAQKGRTQLCALLLAHGADPTLKNQEGQTPLDLVSADDVSALLTAAMPPSA
                     LPSCYKPQVLNGVRSPGATADALSSGPSSPSSLSAASSLDNLSGSFSELSSVVSSSGT
                     EGASSLEKKEVPGVDFSITQFVRNLGLEHLMDIFEREQITLDVLVEMGHKELKEIGIN
                     AYGHRHKLIKGVERLISGQQGLNPYLTLNTSGSGTILIDLSPDDKEFQSVEEEMQSTV
                     REHRDGGHAGGIFNRYNILKIQKVCNKKLWERYTHRRKEVSEENHNHANERMLFHGSP
                     FVNAIIHKGFDERHAYIGGMFGAGIYFAENSSKSNQYVYGIGGGTGCPVHKDRSCYIC
                     HRQLLFCRVTLGKSFLQFSAMKMAHSPPGHHSVTGRPSVNGLALAEYVIYRGEQAYPE
                     YLITYQIMRPEGMVDG"
ORIGIN

  Query Match           62.1%;  Score 2357.8;  DB 6;  Length 3799;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1434 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 1493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 60

Qy       1494 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 1553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 120

Qy       1554 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 1613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 180

Qy       1614 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 1673
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 240

Qy       1674 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1733
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 300

Qy       1734 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 1793
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 360

Qy       1794 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 1853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 420

Qy       1854 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 1913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 480

Qy       1914 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 1973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 540

Qy       1974 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 2033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 600

Qy       2034 TTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGG 2093
             ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        601 TTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGG 660

Qy       2094 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 2153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 720

Qy       2154 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 2213
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 780

Qy       2214 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 2273
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 840

Qy       2274 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2333
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 900

Qy       2334 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2393
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 960

Qy       2394 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2453
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 1020

Qy       2454 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2513
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 1080

Qy       2514 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2573
             |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1081 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 1140

Qy       2574 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2633
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 1200

Qy       2634 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 1260

Qy       2694 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 1320

Qy       2754 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 1380

Qy       2814 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 1440

Qy       2874 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 1500
```

```
Qy       2934 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 2993
Db       1501 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 1560

Qy       2994 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 3053
Db       1561 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 1620

Qy       3054 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 3113
Db       1621 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 1680

Qy       3114 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 3173
Db       1681 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 1740

Qy       3174 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 3233
Db       1741 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 1800

Qy       3234 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 3293
Db       1801 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 1860

Qy       3294 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 3353
Db       1861 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 1920

Qy       3354 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 3413
Db       1921 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 1980

Qy       3414 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 3473
Db       1981 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 2040

Qy       3474 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 3533
Db       2041 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 2100

Qy       3534 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 3593
Db       2101 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 2160

Qy       3594 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 3653
Db       2161 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 2220

Qy       3654 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 3713
Db       2221 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 2280

Qy       3714 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 3773
Db       2281 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 2340

Qy       3774 TACTCCTTTGCTGAAAAAAAA 3794
Db       2341 TACTCCTTTGCTGAAAAAAAA 2361
```

```
RESULT 31
HSM804989
LOCUS       HSM804989              4130 bp    mRNA    linear   PRI 10-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp667A1222 (from clone DKFZp667A1222).
ACCESSION   AL833676
VERSION     AL833676.1  GI:21734324
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4130)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp667A1222) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .4130
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp667A1222"
                     /db_xref="taxon:9606"
                     /clone="DKFZp667A1222"
                     /tissue_type="lymph node"
                     /clone_lib="667 (synonym: hlno2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     polyA_signal    4094. .4099
     polyA_site      4113
ORIGIN
```

```
  Query Match           50.6%;  Score 1922.4;  DB 9;  Length 4130;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1923;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1871 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 1930
Db          1 GATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCA 60

Qy       1931 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 1990
Db         61 GAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCT 120

Qy       1991 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 2050
Db        121 TTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGT 180

Qy       2051 GCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 2110
Db        181 GCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGA 240

Qy       2111 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 2170
Db        241 GATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAG 300

Qy       2171 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2230
Db        301 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 360

Qy       2231 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2290
Db        361 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 420

Qy       2291 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2350
Db        421 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 480

Qy       2351 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 2410
Db        481 TTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAAT 540

Qy       2411 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 2470
Db        541 GCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAG 600

Qy       2471 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 2530
```

```
Db        601 GGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAAT 660
Qy       2531 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 2590
Db        661 CAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACA 720
Qy       2591 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 2650
Db        721 GCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTG 780
Qy       2651 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 2710
Db        781 AGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTT 840
Qy       2711 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 2770
Db        841 TCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTT 900
Qy       2771 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 2830
Db        901 AGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGAT 960
Qy       2831 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 2890
Db        961 TTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAG 1020
Qy       2891 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 2950
Db       1021 AGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATT 1080
Qy       2951 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 3010
Db       1081 GGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCC 1140
Qy       3011 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 3070
Db       1141 GGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTT 1200
Qy       3071 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 3130
Db       1201 ATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACA 1260
Qy       3131 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 3190
Db       1261 GTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTC 1320
Qy       3191 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 3250
Db       1321 AAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAA 1380
Qy       3251 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3310
Db       1381 GTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 1440
Qy       3311 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3370
Db       1441 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 1500
Qy       3371 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3430
Db       1501 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 1560
Qy       3431 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3490
Db       1561 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 1620
Qy       3491 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3550
Db       1621 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 1680
Qy       3551 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3610
Db       1681 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 1740
Qy       3611 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3670
Db       1741 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 1800
Qy       3671 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3730
Db       1801 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 1860
Qy       3731 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 3790
Db       1861 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAA 1920
Qy       3791 AAAA 3794
Db       1921 AAAA 1924
```

RESULT 32
BC057370

LOCUS	BC057370 4054 bp mRNA linear ROD 08-OCT-2003
DEFINITION	Mus musculus tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase, mRNA (cDNA clone MGC:67299 IMAGE:6835784), complete cds.
ACCESSION	BC057370
VERSION	BC057370.1 GI:34980998
KEYWORDS	MGC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 4054)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE	22388257
PUBMED	12477932
REFERENCE	2 (bases 1 to 4054)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (02-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 126 Row: p Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES Location/Qualifiers

source 1. .4054
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:67299 IMAGE:6835784"
/tissue_type="Brain, enriched mouse brain 12.5dpc"
/clone_lib="NIH_BMAP_FI0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

gene 1. .4054
/gene="Tnks"
/db_xref="LocusID:21951"
/db_xref="MGI:1341087"

CDS 7. .3969
/codon_start=1
/product="Tnks protein"
/protein_id="AAH57370.1"
/db_xref="GI:34980999"
/db_xref="LocusID:21951"
/translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPPLSPGLAPGPT
PASPTAGGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGAVCTVAAPAAVPAASA
AVGVAPTPAGGGGGGNNSASSASSPTSSSSSSPSSPGSSLAESPEAAGVGSTATLGA
GAAGLGPGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSSPLHFAAG
FGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARDNWNY
TPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYKKDELLE
AARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADVHAKD
KGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSLLLSH
GADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLALEIIN
FKQPQSHETALHCAVASLHPKRKQVAELLLRKGANVNEKNKDFMTPLHVAAERAHNDV
MEVLHKHGAKMNALDSLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTAAQMG
NEAVQQILSESTPMRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHSTPLH
FAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASVNVAD
LWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLLRGDA
ALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEHGADV
NAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQLCAL
LLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVVSASLISP
ASTPSCLSAASSIDNLTGPLTDLAVGGASNAGDGAAGAERKEGEVAGLDMNISQFLKS
LGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQGTNP
YLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIRIQKV
VNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGMFGAG
IYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFSTMKMA
HAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTATAAEQKT
"

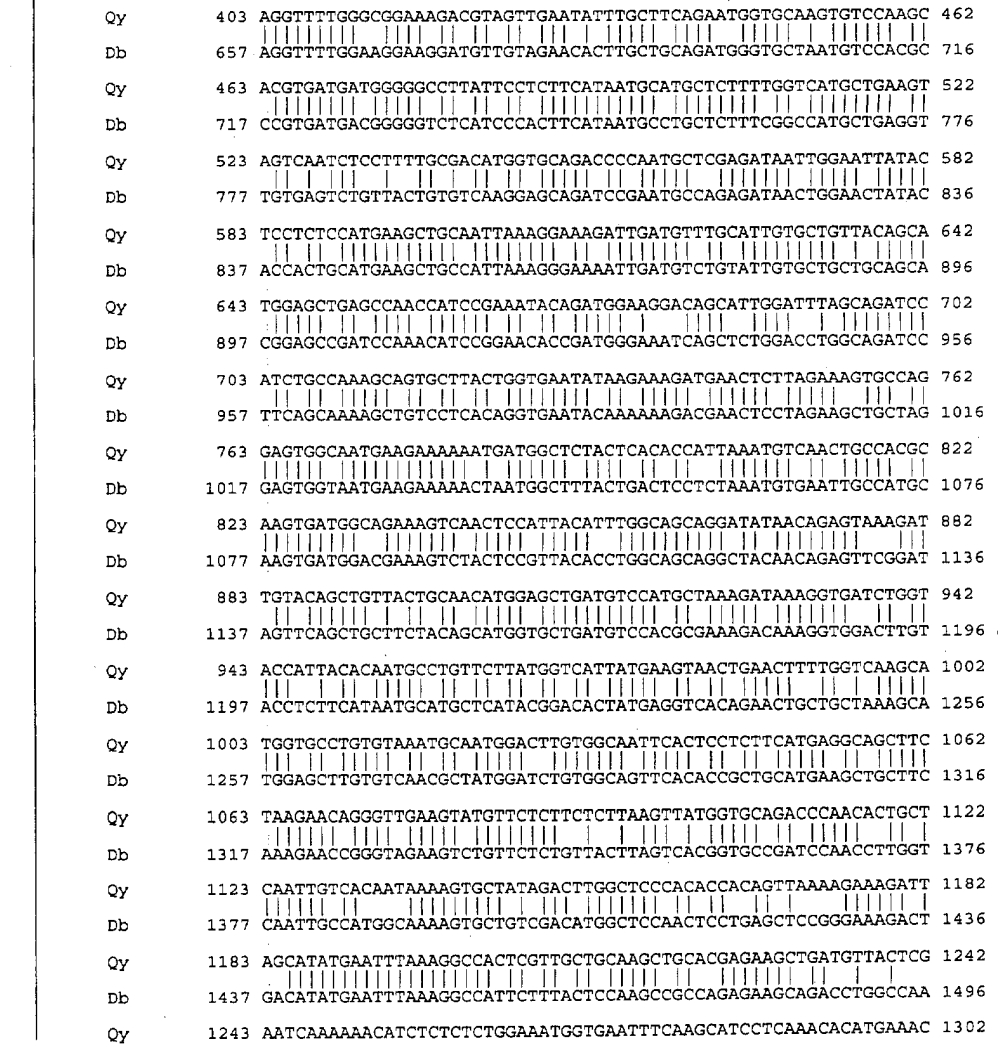
misc_feature 538. .1032
/note="Arp; Region: FOG: Ankyrin repeat [General function prediction only]"
/db_xref="CDD:COG0666"

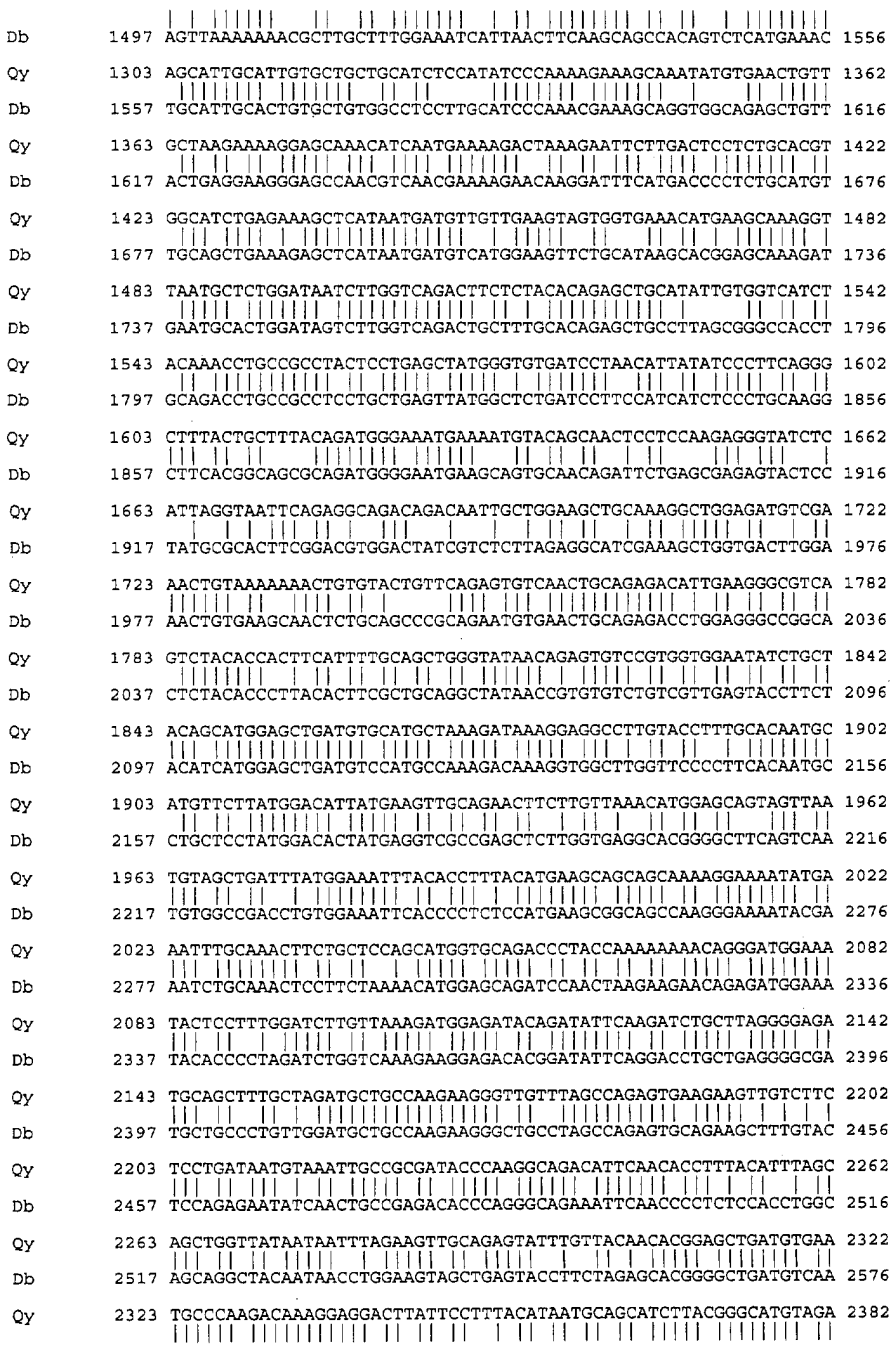
misc_feature 3073. .3246
/note="SAM; Region: Sterile alpha motif. Widespread domain in signalling and nuclear proteins. In EPH-related tyrosine kinases, appears to mediate cell-cell initiated signal transduction via the binding of SH2-containing proteins to a conserved tyrosine that is phosphorylated. In many cases mediates homodimerisation"
/db_xref="CDD:smart00454"

misc_feature 3424. .3903
/note="PARP; Region: Poly(ADP-ribose) polymerase catalytic domain. Poly(ADP-ribose) polymerase catalyses the covalent attachment of ADP-ribose units from NAD+ to itself and to a limited number of other DNA binding proteins, which decreases their affinity for DNA. Poly(ADP-ribose) polymerase is a regulatory component induced by DNA damage. The carboxyl-terminal region is the most highly conserved region of the protein. Experiments have shown that a carboxyl 40 kDa fragment is still catalytically active"
/db_xref="CDD:pfam00644"

ORIGIN

Query Match 48.5%; Score 1839.8; DB 10; Length 4054;
Best Local Similarity 73.0%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 862; Indels 27; Gaps 2;




Db 1497 AGTTAAAAAAACGCTTGCTTTGGAAATCATTAACTTCAAGCAGCCACAGTCTCATGAAAC 1556
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1557 TGCATTGCACTGTGCTGTGGCCTCCTTGCATCCCAAACGAAAGCAGGTGGCAGAGCTGTT 1616
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1617 ACTGAGGAAGGGAGCCAACGTCAACGAAAAGAACAAGGATTTCATGACCCCTCTGCATGT 1676
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1677 TGCAGCTGAAAGAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCACGGAGCAAAGAT 1736
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1737 GAATGCACTGGATAGTCTTGGTCAGACTGCTTTGCACAGAGCTGCCTTAGCGGGCCACCT 1796
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1797 GCAGACCTGCCGCCTCCTGCTGAGTTATGGCTCTGATCCTTCCATCATCTCCCTGCAAGG 1856
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1857 CTTCACGGCAGCGCAGATGGGGAATGAAGCAGTGCAACAGATTCTGAGCGAGAGTACTCC 1916
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1917 TATGCGCACTTCGGACGTGGACTATCGTCTCTTAGAGGCATCGAAAGCTGGTGACTTGGA 1976
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1977 AACTGTGAAGCAACTCTGCAGCCCGCAGAATGTGAACTGCAGAGACCTGGAGGGCCGGCA 2036
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 2037 CTCTACACCCTTACACTTCGCTGCAGGCTATAACCGTGTGTCTGTCGTTGAGTACCTTCT 2096
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 2097 ACATCATGGAGCTGATGTCCATGCCAAAGACAAAGGTGGCTTGGTTCCCCTTCACAATGC 2156
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 2157 CTGCTCCTATGGACACTATGAGGTCGCCGAGCTCTTGGTGAGGCACGGGGCTTCAGTCAA 2216
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2217 TGTGGCCGACCTGTGGAAATTCACCCCTCTCCATGAAGCGGCAGCCAAGGGAAAATACGA 2276
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2277 AATCTGCAAACTCCTTCTAAAACATGGAGCAGATCCAACTAAGAAGAACAGAGATGGAAA 2336
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2337 TACACCCCTAGATCTGGTCAAAGAAGGAGACACGGATATTCAGGACCTGCTGAGGGGCGA 2396
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2397 TGCTGCCCTGTTGGATGCTGCCAAGAAGGGCTGCCTAGCCAGAGTGCAGAAGCTTTGTAC 2456
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2457 TCCAGAGAATATCAACTGCCGAGACACCCAGGGCAGAAATTCAACCCCTCTCCACCTGGC 2516
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2517 AGCAGGCTACAATAACCTGGAAGTAGCTGAGTACCTTCTAGAGCACGGGGCTGATGTCAA 2576
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2577 TGCCCAGGACAAAGGAGGGCTAATCCCACTTCACAACGCTGCGTCTTATGGGCATGTTGA 2636
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2637 CATAGCAGCGTTACTGATAAAATACAACACGTGTGTAAACGCAACGGATAAGTGGGCCTT 2696
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2697 TACTCCCCTTCACGAAGCTGCCCAGAAAGGAAGGACGCAGTTATGTGCCCTGCTCCTAGC 2756
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2757 ACATGGCGCAGACCCAACCATGAAGAACCAGGAAGGTCAGACACCTTTAGATCTGGCGAC 2816
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2817 GGCTGATGACATCCGGGCTCTGCTGATAGATGCCATGCCCCCTGAGGCCTTACCTACCTG 2876
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2877 TTTTAAACCTCAGG----------CAACTGTAGTTAGTGCCTCTCTGAT 2915
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2916 CTCACCAGCGTCCACCCCCTCCTGCCTCTCTGCTGCCAGCAGCATAGATAACCTCACTGG 2975
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2976 CCCGTTAACGGATCTGGCTGTAGGAGGGGCCTCCAATGCAGGGGATGGCGCTGCGGGTGC 3035
Qy 2803 GGAGAAAAAG------GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 2856
Db 3036 AGAAAGGAAGGAAGGAGAAGTTGCGGGTCTTGACATGAACATCAGCCAGTTCCTTAAGAG 3095
Qy 2857 TCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATT 2916
Db 3096 CCTTGGCCTTGAGCACCTGAGGGACATTTTTGAAACAGAGCAGATTACGCTAGATGTGTT 3155
Qy 2917 AGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCA 2976
Db 3156 GGCTGATATGGGTCATGAAGAATTGAAAGAAATAGGCATCAATGCATATGGACACCGCCA 3215
Qy 2977 CAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTT 3036
Db 3216 CAAGTTGATCAAGGGTGTTGAAAGACTTTTAGGTGGACAACAAGGAACCAATCCATATTT 3275
Qy 3037 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3096
Db 3276 GACTTTTCACTGTGTGAATCAGGGAACTATTTTGCTAGATCTTGCTCCAGAAGATAAAGA 3335
Qy 3097 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3156
Db 3336 ATATCAATCAGTAGAAGAAGAGATGCAGAGCACTATTCGAGAGCACAGAGACGGTGGGAA 3395
Qy 3157 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 3216
Db 3396 CGCTGGCGGCATCTTCAACAGATACAACGTCATTCGAATTCAGAAAGTTGTGAATAAGAA 3455
Qy 3217 ACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGC 3276
Db 3456 GCTGAGGGAGCGATTCTGCCACAGACAAAAGGAGGTGTCTGAGGAGAACCACAACCATCA 3515
Qy 3277 CAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTT 3336
Db 3516 CAATGAGCGCATGCTGTTCCACGGTTCTCCTTTCATTAATGCAATTATTCATAAAGGGTT 3575
Qy 3337 TGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAA 3396
Db 3576 TGATGAGCGACATGCGTACATTGGAGGGATGTTTGGAGCTGGAATTTACTTTGCTGAAAA 3635
Qy 3397 CTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCA 3456
Db 3636 CTCCTCAAAAAGCAATCAGTATGTTTATGGAATTGGAGGAGGAACAGGCTGTCCCACACA 3695

```
Qy       3457 CAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGG 3516
Db       3696 CAAAGACAGATCATGTTATATATGTCACAGACAAATGCTTTTCTGTAGAGTGACCCTTGG 3755

Qy       3517 AAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTC 3576
Db       3756 AAAGTCCTTCCTGCAATTCAGCACCATGAAAATGGCCCATGCACCCCCAGGCCATCACTC 3815

Qy       3577 AGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGG 3636
Db       3816 AGTCATCGGTAGACCAAGTGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGG 3875

Qy       3637 AGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3689
Db       3876 GGAACAGGCATATCCAGAATATCTCATCACCTACCAGATCATGAAGCCAGAAG 3928
```

RESULT 33
AX062144
LOCUS AX062144 3984 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 3 from Patent WO0100849.
ACCESSION AX062144
VERSION AX062144.1 GI:12540047
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
TITLE Tankyrase2 materials and methods
JOURNAL Patent: WO 0100849-A 3 04-JAN-2001;
ICOS CORPORATION (US)
FEATURES Location/Qualifiers
source 1. .3984
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 1. .3984
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC25053.1"
/db_xref="GI:12540048"
/db_xref="REMTREMBL:CAC25053"
/translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPPLSPGLAPGTT
PASPTASGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGTSCCSTTSTICTVAAA
PVVPAVSTSSAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPGSSLAESPEAAGVS
STAPLGPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSS
PLHFAAGFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPN
ARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEY
KKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHG
ADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEV
CSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKT
LALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQG
FTAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEG
RHSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHG
ASVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQ
DLLKGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYL
LEHGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKG
RTQLCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVV
SASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNAGDGAAGTERKEGEVAGLDMN
ISQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLG
GQQGTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYN
VIRIQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYI
GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQ
FSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTA
TAAEQKT"
ORIGIN

Query Match 47.8%; Score 1813.4; DB 6; Length 3984;
Best Local Similarity 72.3%; Pred. No. 0;

Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        652 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 711

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        712 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 771

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        772 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 831

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        832 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 891

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        892 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 951

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        952 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1011

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1012 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1071

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1072 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1131

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1132 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1191

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db       1192 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1251

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db       1252 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1311

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db       1312 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1371

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db       1372 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1431

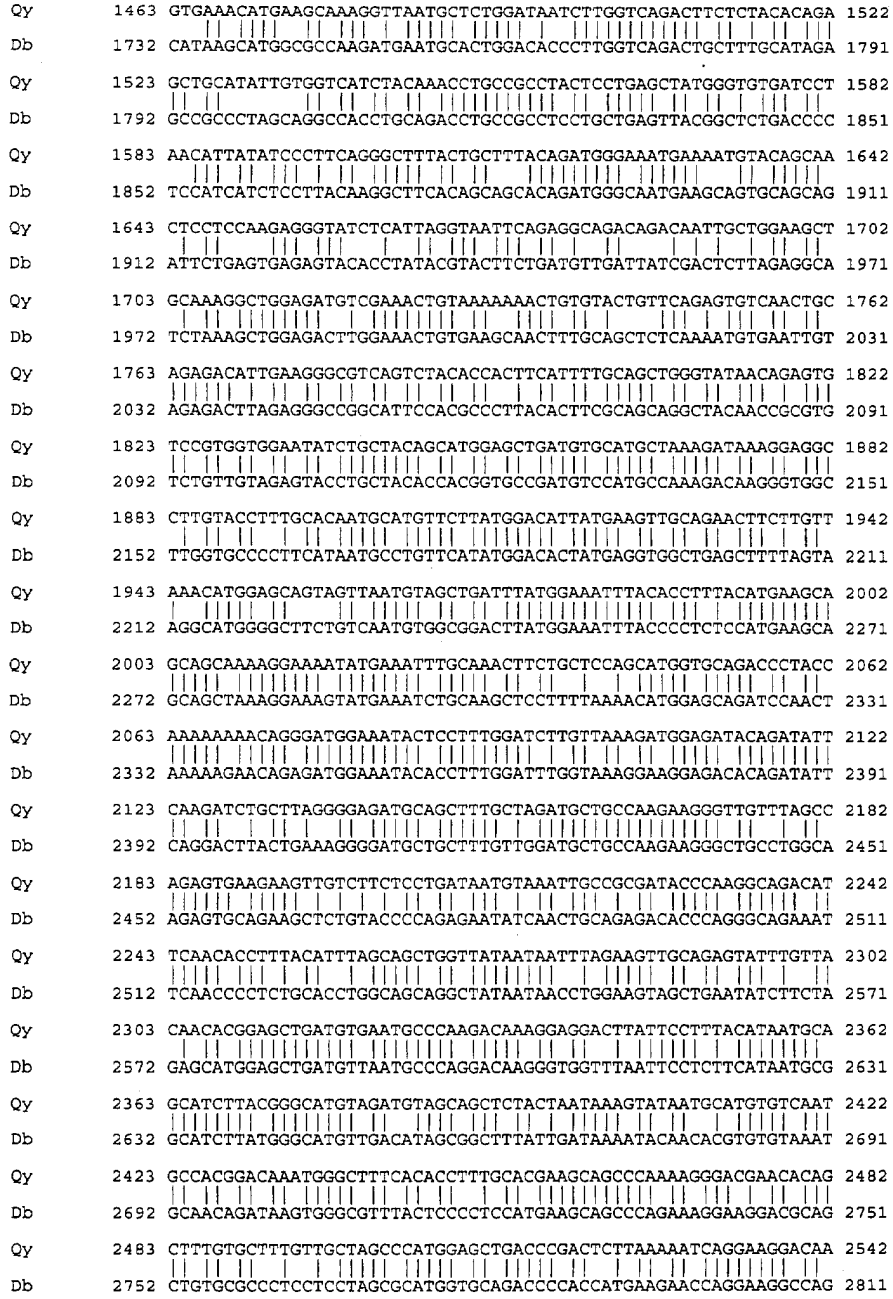
Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db       1432 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1491

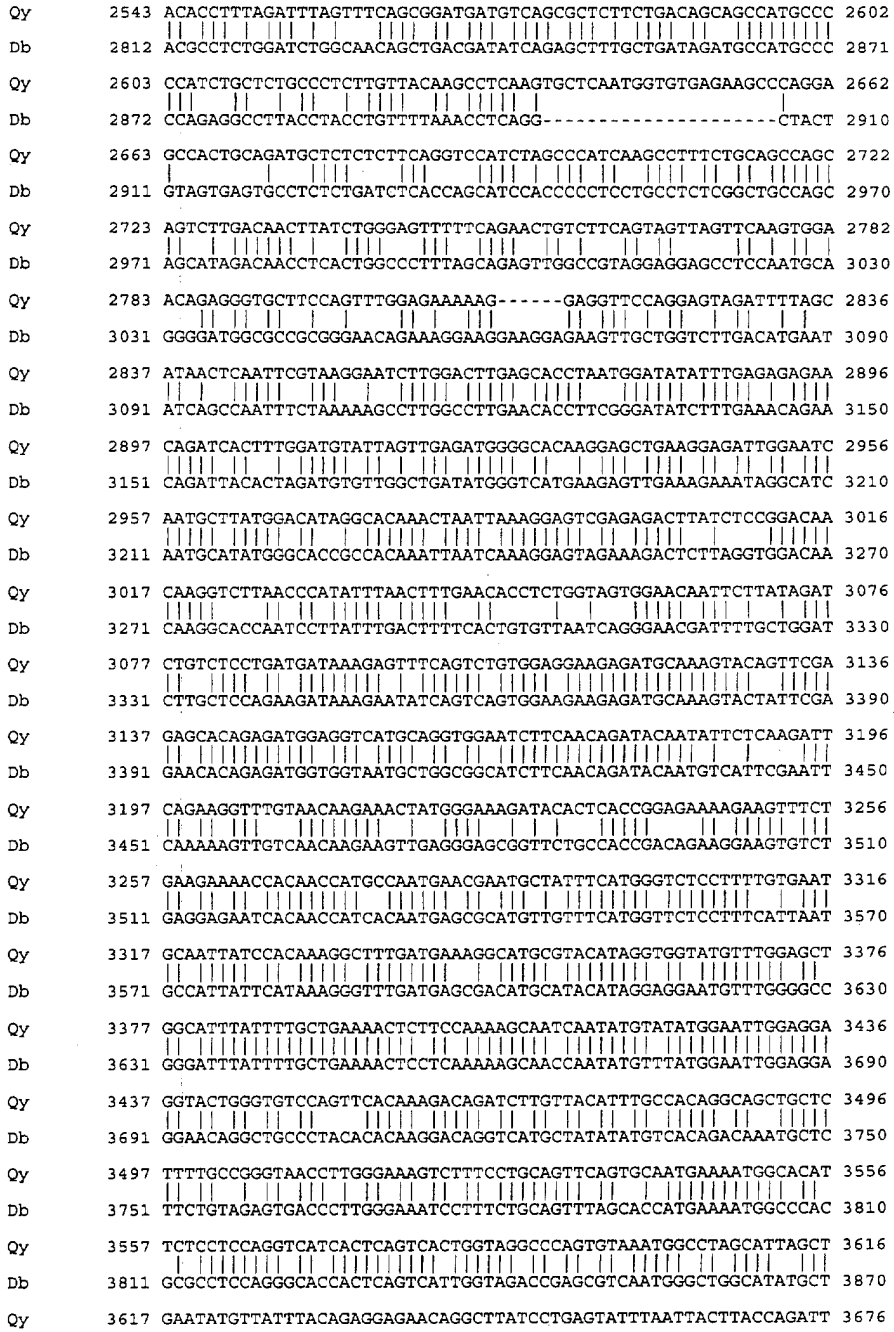
Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db       1492 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1551

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db       1552 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1611

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db       1612 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1671

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db       1672 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1731
```





```
                ||||||||| || ||||||||||||||||| || || |||||| | || ||||||||||
Db       3871 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3930

Qy       3677 ATGAGGCCTGAAG 3689
                ||||  |||  ||||
Db       3931 ATGAAGCCAGAAG 3943


RESULT 34
BD235110
LOCUS       BD235110                4134 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Tankyrase, TRFI-binding protein.
ACCESSION   BD235110
VERSION     BD235110.1  GI:33044880
KEYWORDS    JP 2002517251-A/1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4134)
  AUTHORS   Lange,T.D. and Smith,S.
  TITLE     Tankyrase, TRFI-binding protein
  JOURNAL   Patent: JP 2002517251-A 1 18-JUN-2002;
            THE ROCKEFELLER UNIVERSITY
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002517251-A/1
            PD   18-JUN-2002
            PF   09-JUN-1999 JP 2000553596
            PR   10-JUN-1998 US    09/095225,17-AUG-1998 US    09/135233 PR
            19-NOV-1998 US    09/196387
            PI   TITIA DE LANGE,SUSAN SMITH
            PC   C12N15/00,A01K67/027,A61K31/166,A61K38/45,A61K39/395,A61K39/
            PC   395,A61K45/00,
            PC   A61K48/00,A61P35/00,A61P43/00,C07K16/40,C12N9/12,C12Q1/48, PC
             G01N33/15,
            PC   G01N33/50,C12N15/00,A61K37/52
            CC   Tankyrase, TRFI-binding protein
            FH   Key             Location/Qualifiers
            FT   source          1. .4134
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .4134
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             47.8%;  Score 1813.4;  DB 6;  Length 4134;
  Best Local Similarity   72.3%;  Pred. No. 0;
  Matches 2395;  Conservative    0;  Mismatches  891;  Indels   27;  Gaps    2;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
              || | |  ||  |   |  |||||||||  |||| || || || ||| | || || |||
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
              |  |||||  | ||||||  || ||||||||||| || || || || ||||||||||| ||
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
              |||||||| |||||||| || || | |||  | ||| | || || || || || ||||||
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
              | ||||| |||| |||| |||| |||||||||| |||||||| ||||| |||||
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
              |||||||||||| | ||||| |||||||| |||| ||| || || || ||||| |   ||
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
              ||  ||||  | |||||||| || || ||||| || ||||| |||||||| ||||||||
```

```
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
              |||||| |||||  ||| |||||||| |||||||||||| | |||||| |||| || ||
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
              ||||||| || ||||| |||||||||||  ||||||| ||||| |||||| | |||||
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
              || || ||||||||    || || |||||  | ||| || |||||  |||||||| |||||
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
              ||||| |||||||   || || ||  | || |||||  ||||| |||||||||||||||
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
              || |||||  |  | |||||||| || |||||  |||||  |||||  | |||||  || |||
Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
              || || || |||||  |||||  ||||||  | || |||||  || |||  |  |  || ||
Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
              || || || ||  ||  |  |||| ||  ||     |||||||| | ||  |||||| ||
Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
              ||  ||  |  ||  |||||  | |||||||||||||| || || || || |||||  ||
Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
              |||||||  ||  |  ||  | | ||||||   ||| ||||||||||  | |||||||
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
              || || ||  | |||||||||||  |||| |||||||  || ||||   |||||||| |
Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
              || ||| |   ||| |||| || |||||||||||||   | |||||||| | ||||||
Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
              ||| ||||||| ||||| || || | ||| |||||||||| | |||| ||||||| || ||
Db       1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736

Qy       1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
                 || ||||  || ||| | |||||  |||||  |  |||||||||||  || | ||| |||
Db       1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796

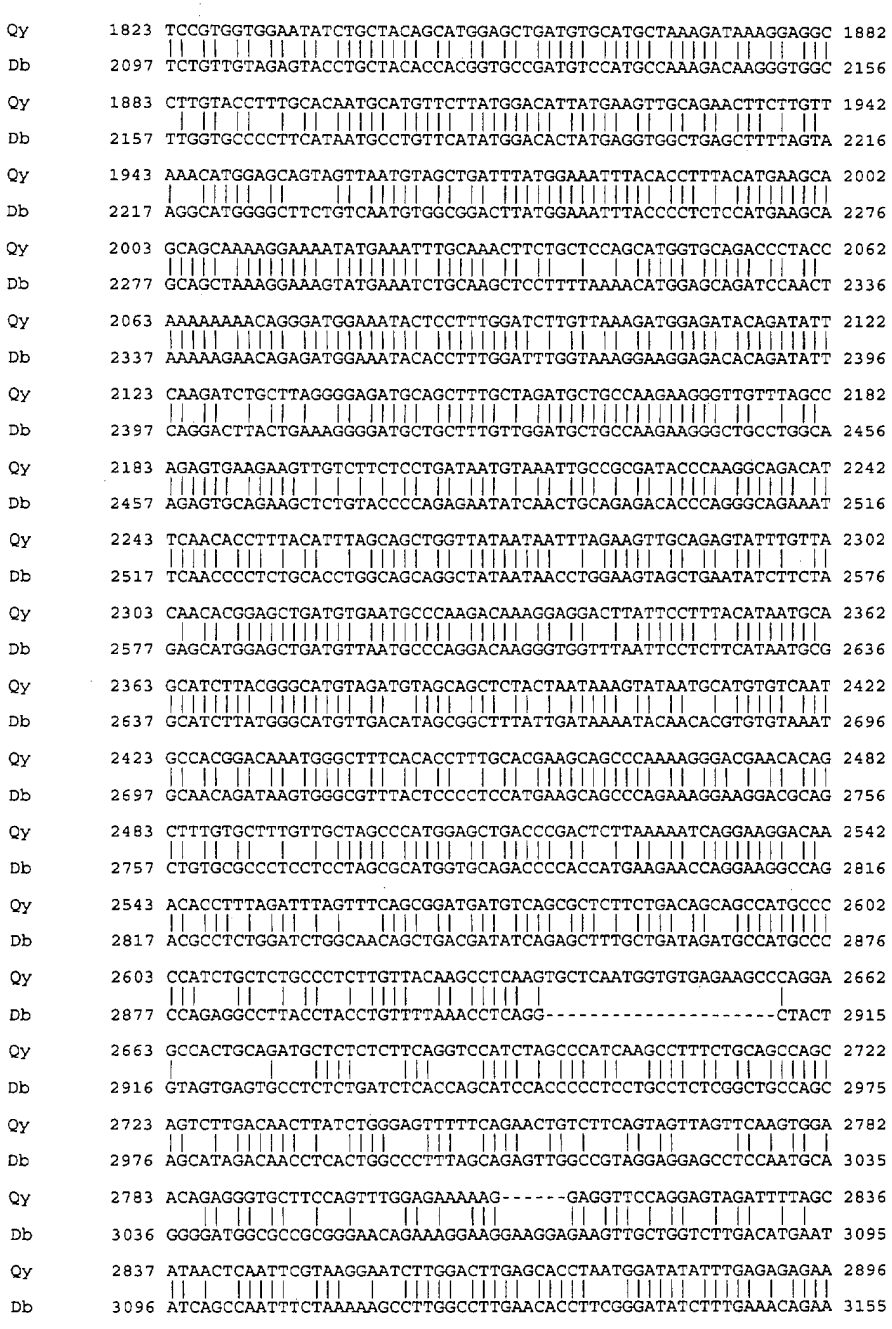
Qy       1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
              || ||          || || || || ||||||||||| || ||||| || | | ||| ||
Db       1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856

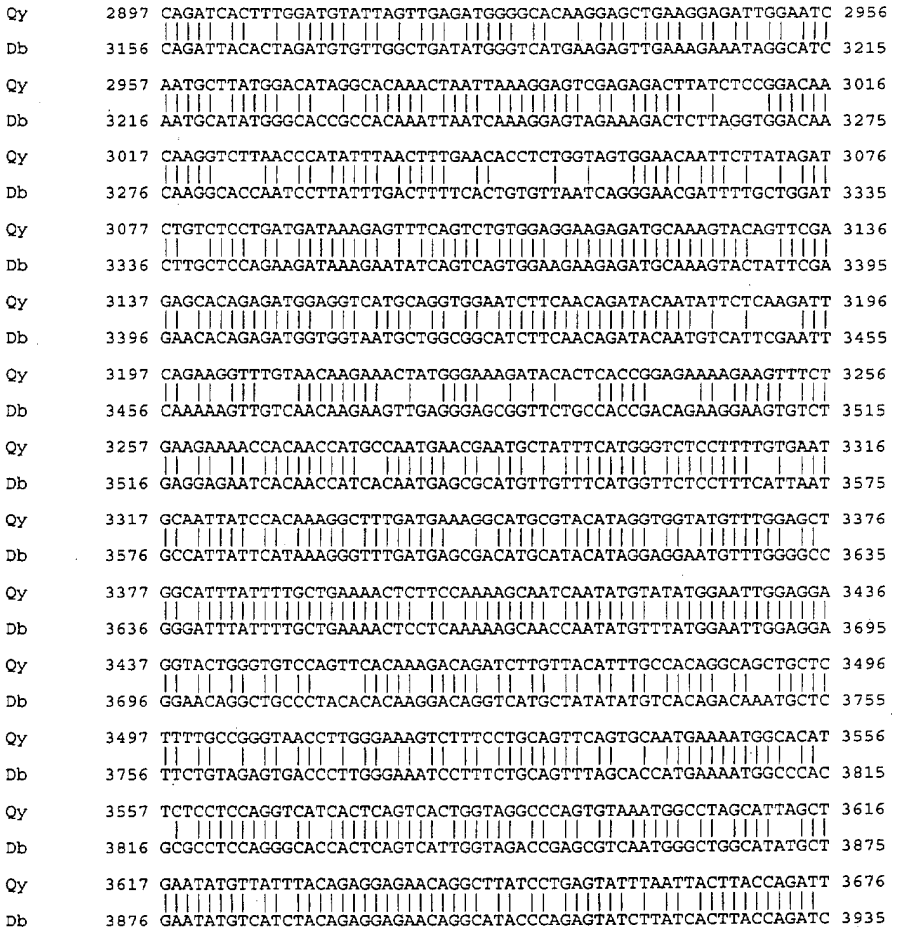
Qy       1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
                 ||| || ||| | || ||||| || ||   ||||||||| ||||||   || |||||
Db       1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916

Qy       1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
              | ||    ||| |||    |  || ||| ||| || |  ||  | | |  | ||| ||
Db       1917 ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCA 1976

Qy       1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
              | || |||||||  | |||||||| ||  |||| || |   ||| | ||| || ||
Db       1977 TCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGT 2036

Qy       1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
              ||||||  | || || || || || || ||  | || || ||||| || || ||| | |||
Db       2037 AGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTG 2096
```

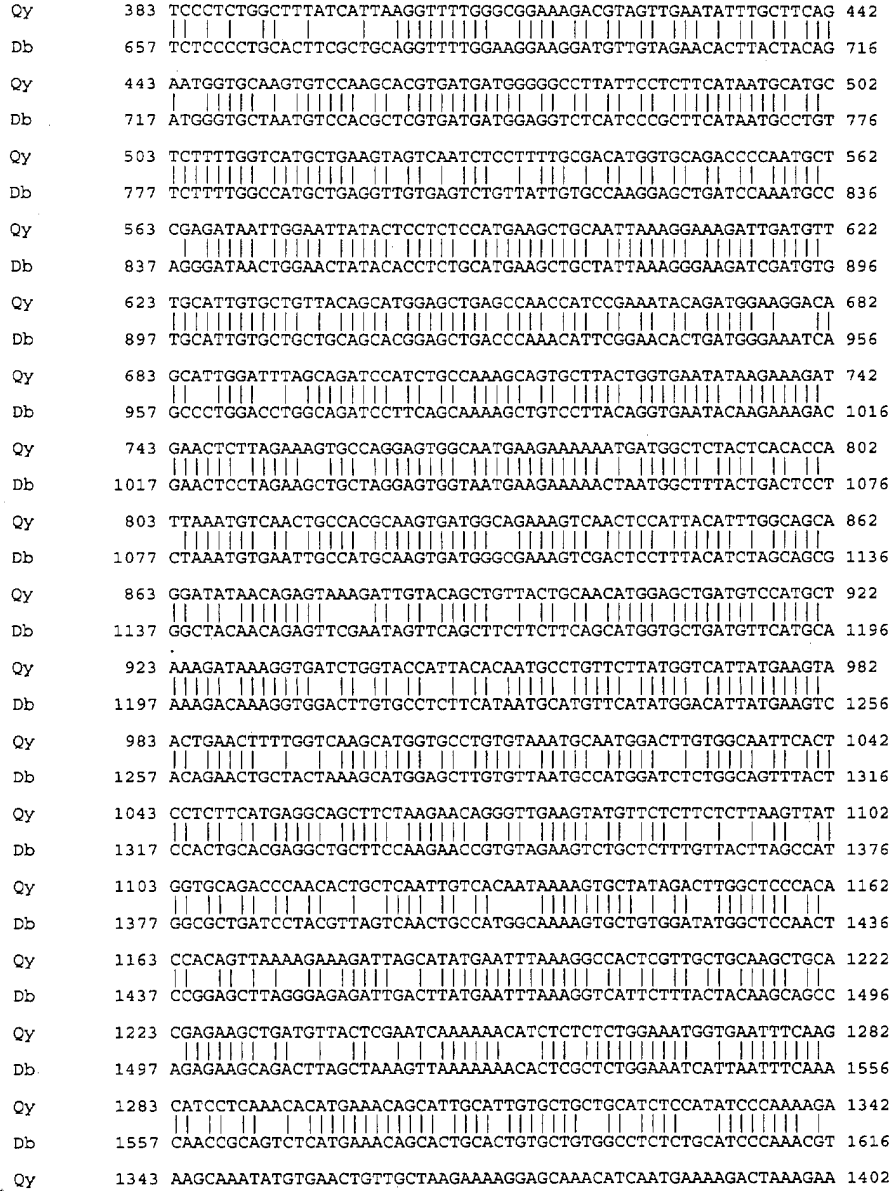


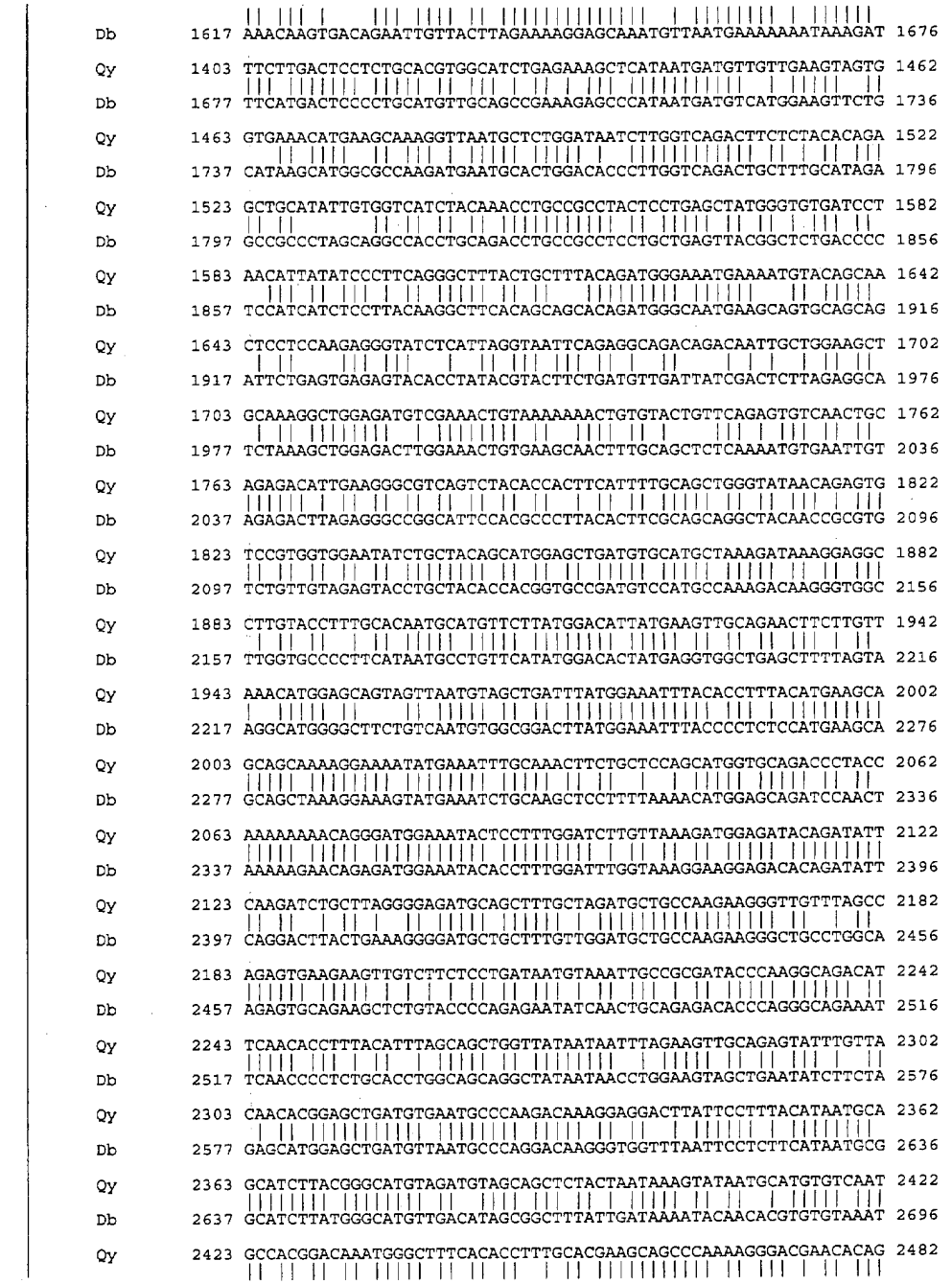


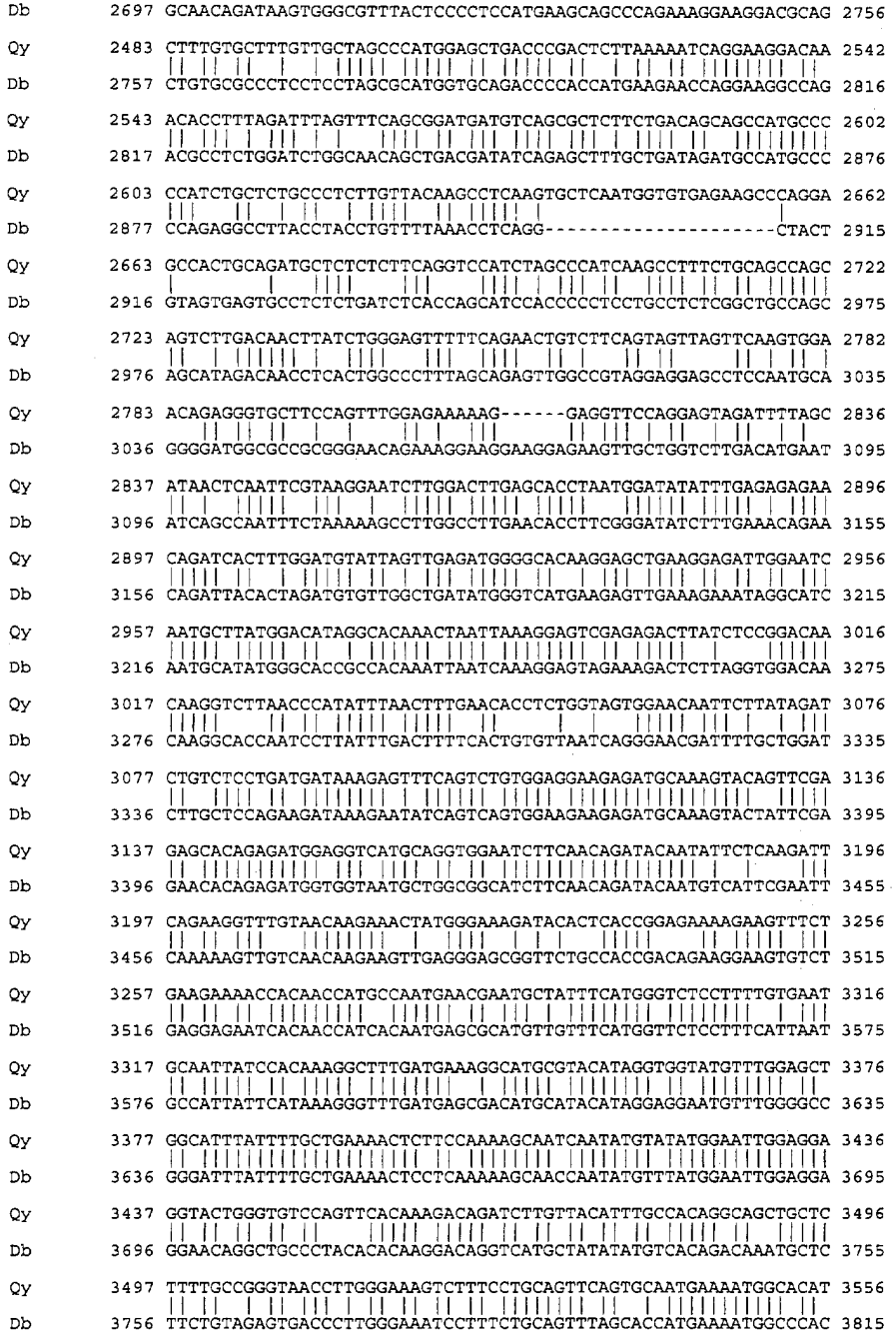
```
RESULT 35
AR274521
LOCUS       AR274521                4134 bp    mRNA    linear   PAT 10-APR-2003
DEFINITION  Sequence 1 from patent US 6506587.
ACCESSION   AR274521
VERSION     AR274521.1  GI:29707021
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4134)
  AUTHORS   De Lange,T. and Smith,S.
  TITLE     TRF 1 binding protein, methods of use thereof
  JOURNAL   Patent: US 6506587-A 1 14-JAN-2003;
FEATURES             Location/Qualifiers
     source          1. .4134
```

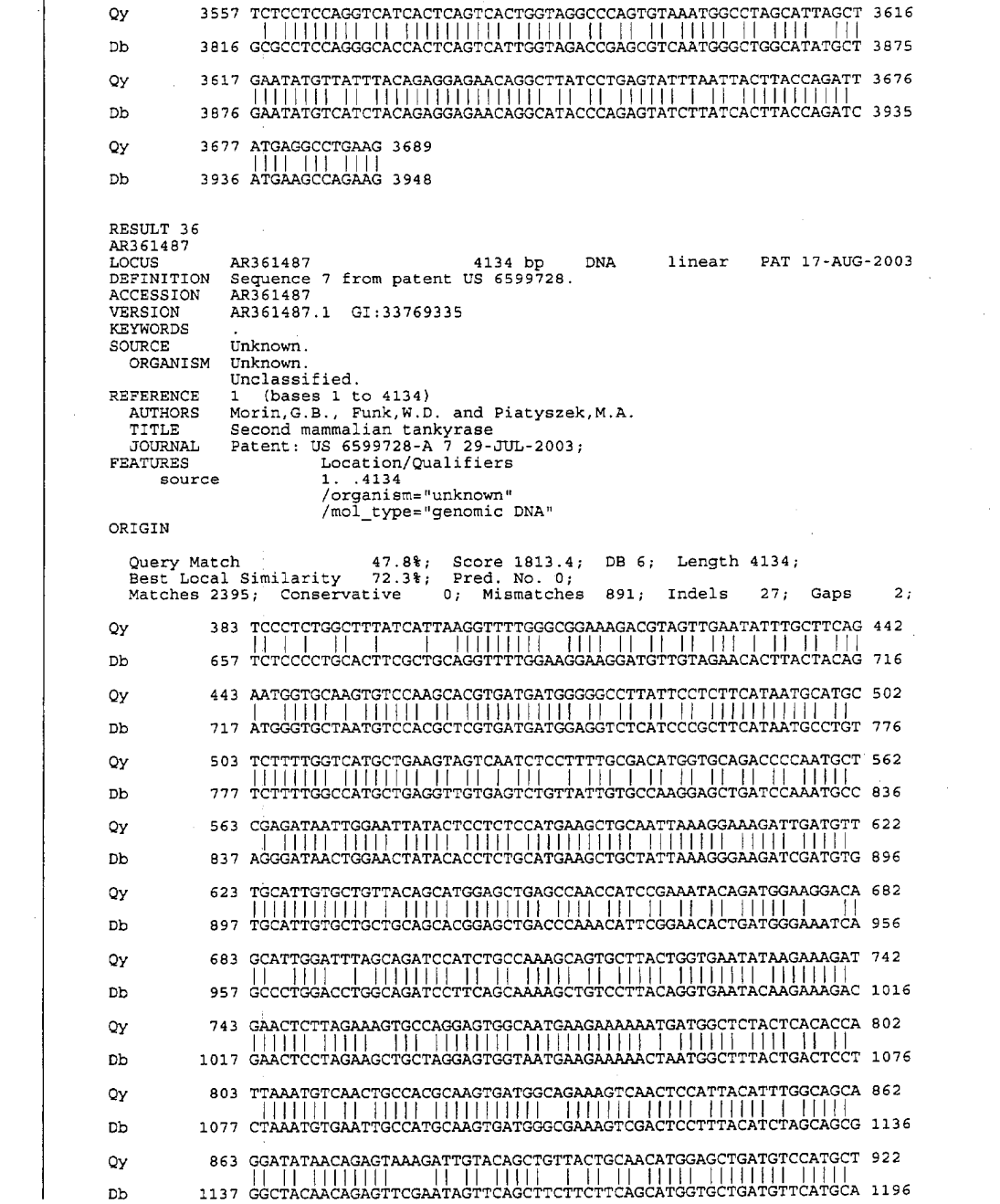
/organism="unknown"
/mol_type="mRNA"
ORIGIN

Query Match 47.8%; Score 1813.4; DB 6; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;



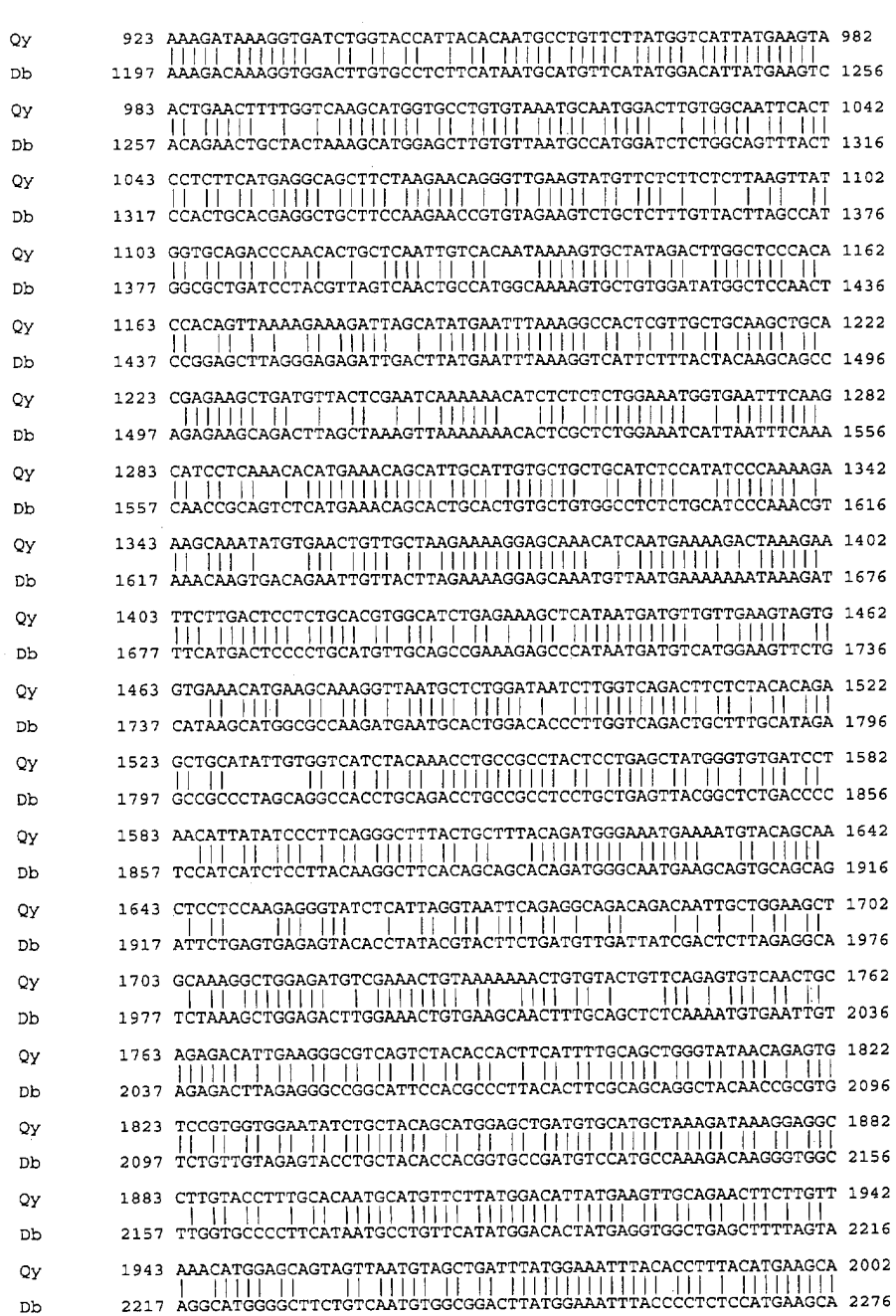


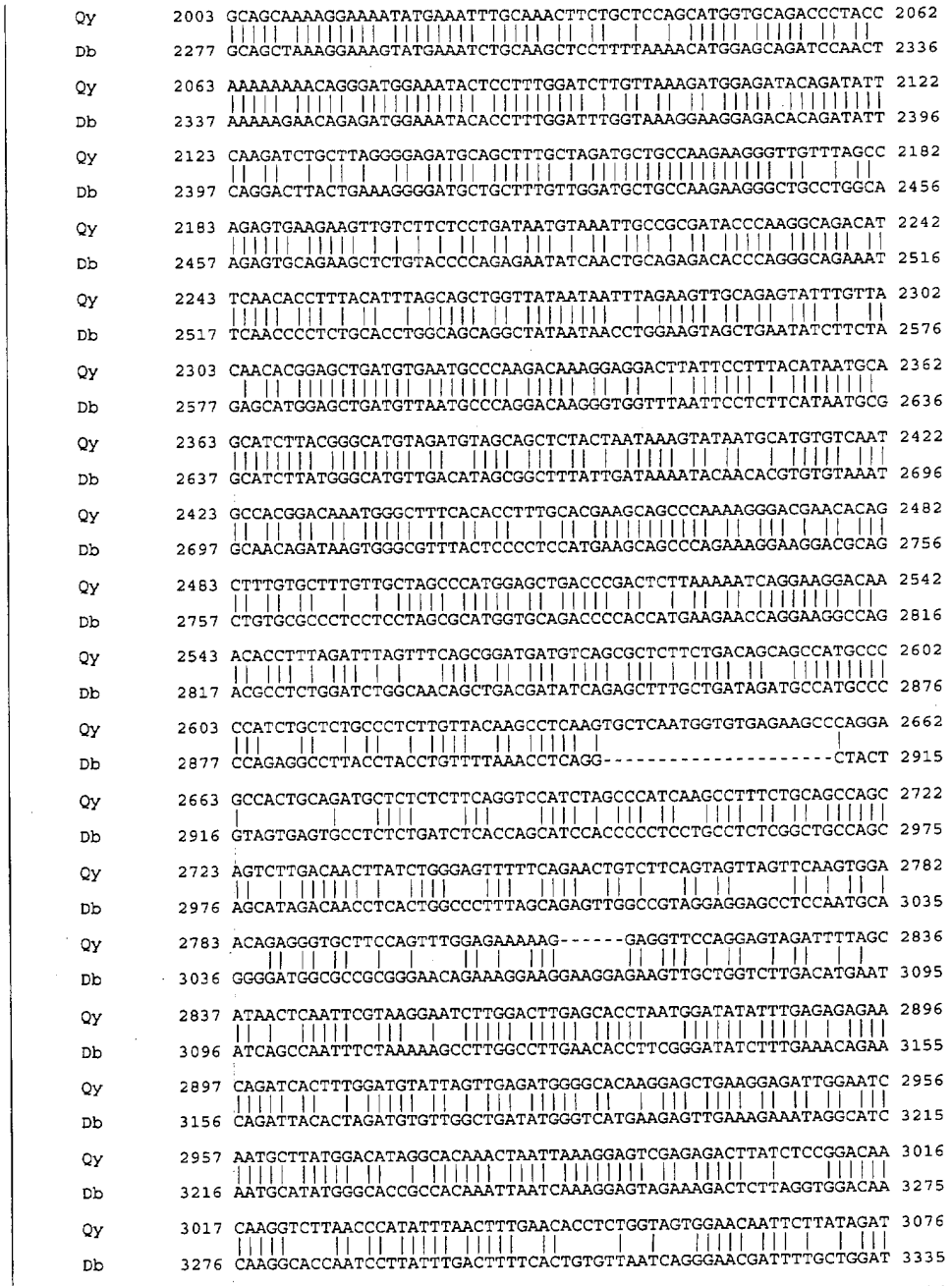


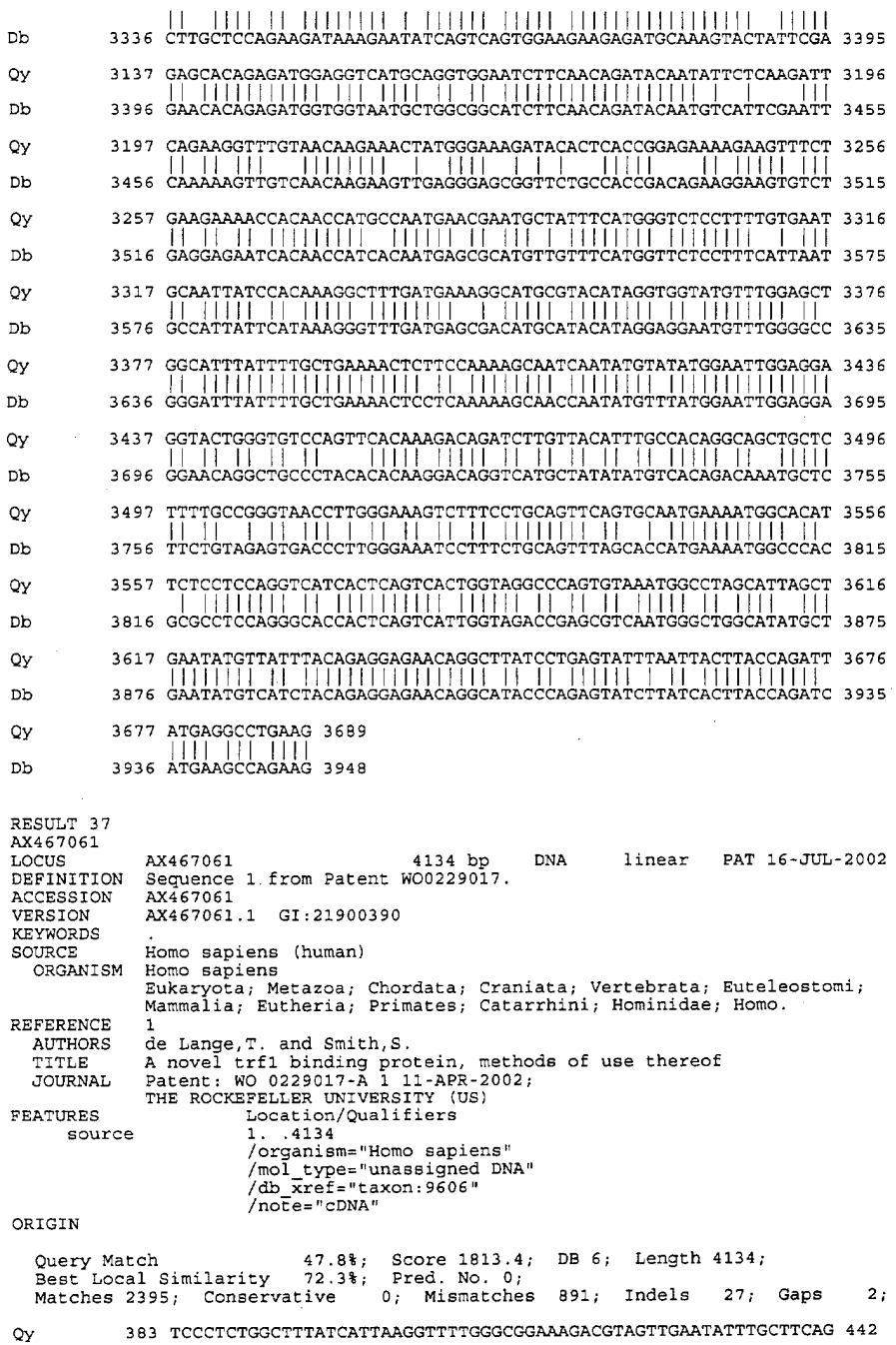


RESULT 36
AR361487
LOCUS AR361487 4134 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 7 from patent US 6599728.
ACCESSION AR361487
VERSION AR361487.1 GI:33769335
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4134)
AUTHORS Morin,G.B., Funk,W.D. and Piatyszek,M.A.
TITLE Second mammalian tankyrase
JOURNAL Patent: US 6599728-A 7 29-JUL-2003;
FEATURES Location/Qualifiers
source 1. .4134
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 47.8%; Score 1813.4; DB 6; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

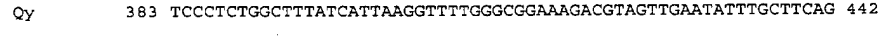


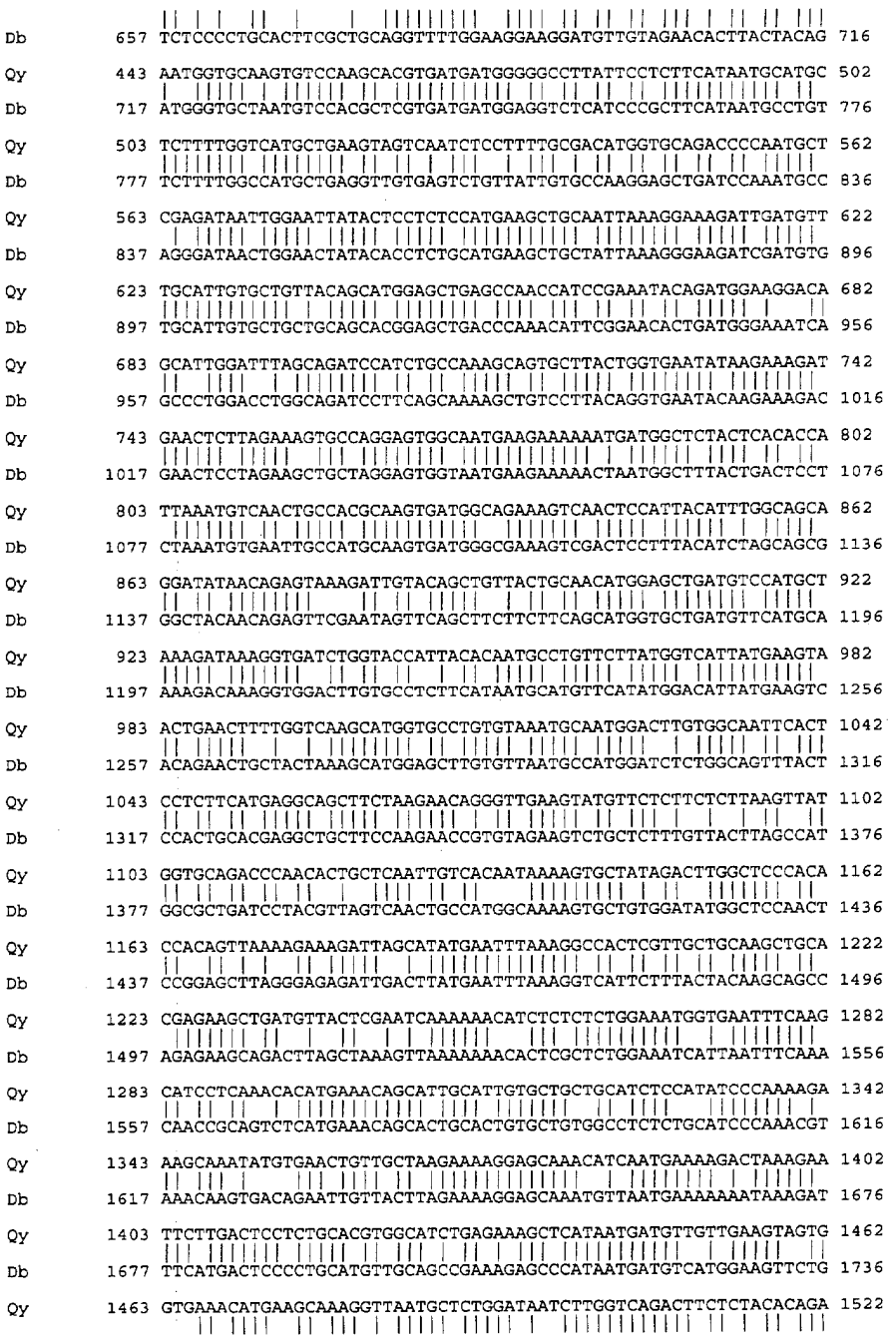


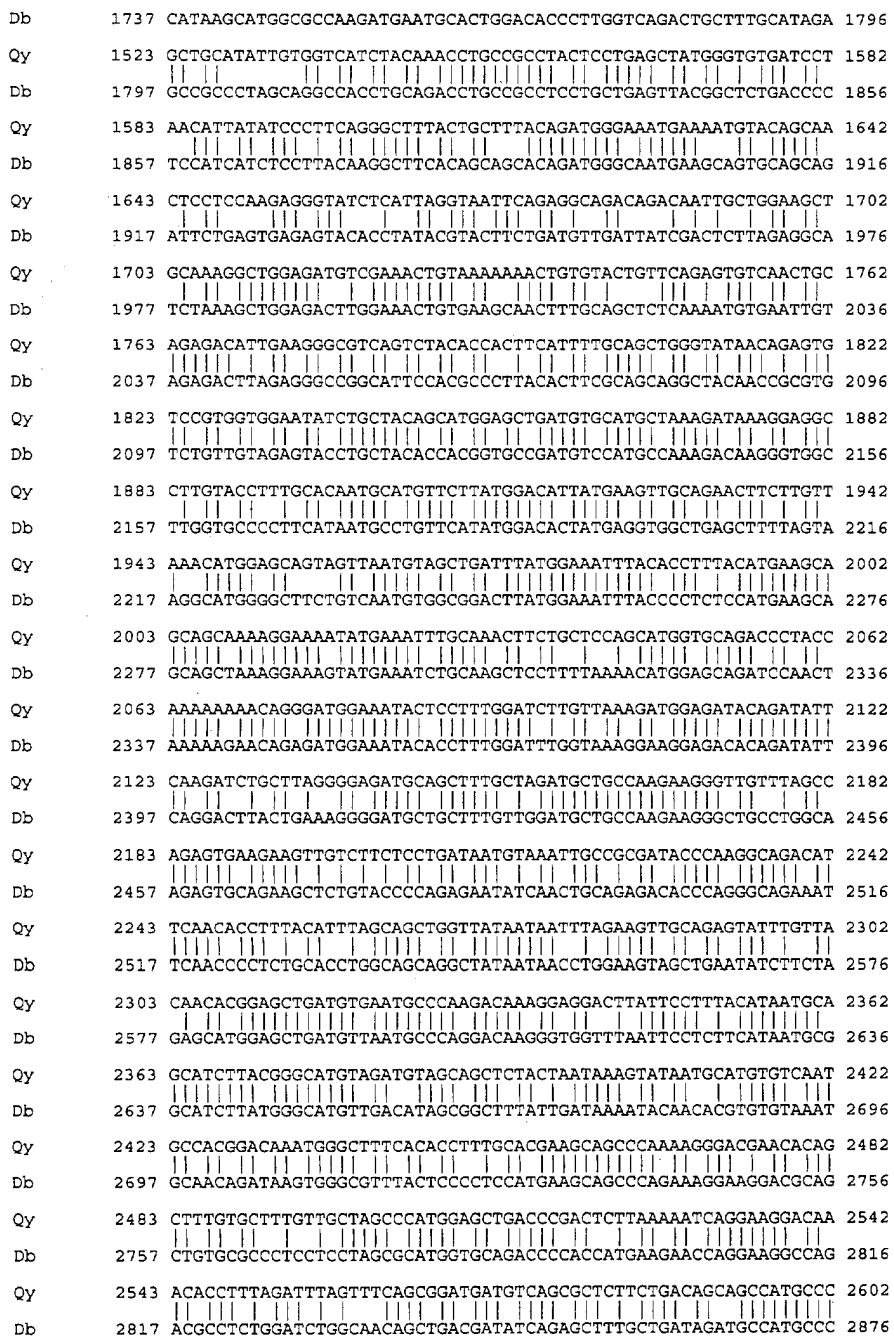


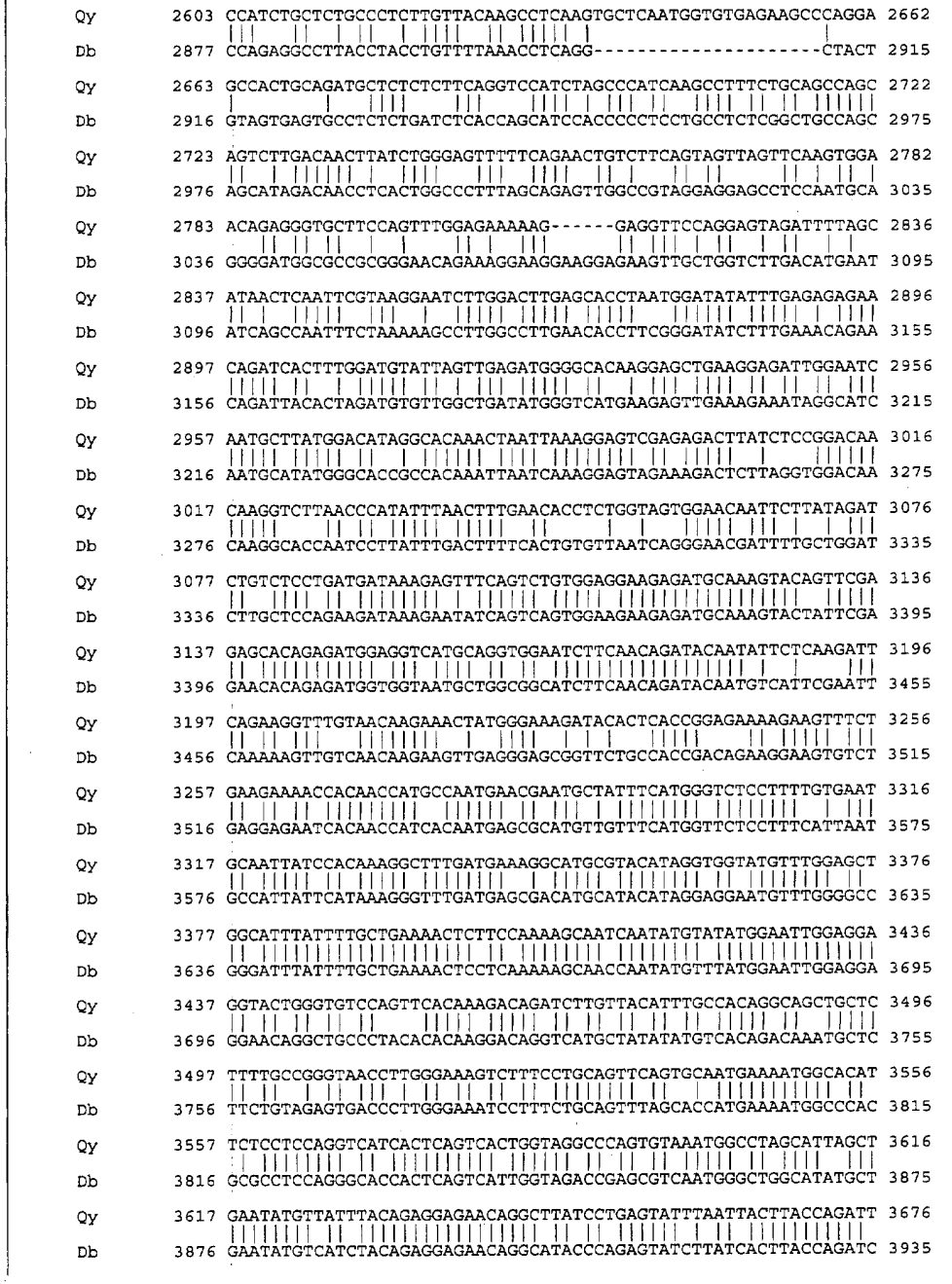
RESULT 37
AX467061
LOCUS AX467061 4134 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO0229017.
ACCESSION AX467061
VERSION AX467061.1 GI:21900390
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS de Lange,T. and Smith,S.
TITLE A novel trf1 binding protein, methods of use thereof
JOURNAL Patent: WO 0229017-A 1 11-APR-2002;
THE ROCKEFELLER UNIVERSITY (US)
FEATURES Location/Qualifiers
source 1. .4134
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="cDNA"
ORIGIN

Query Match 47.8%; Score 1813.4; DB 6; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;









```
Qy       3677 ATGAGGCCTGAAG 3689
              |||| ||| ||||
Db       3936 ATGAAGCCAGAAG 3948
```

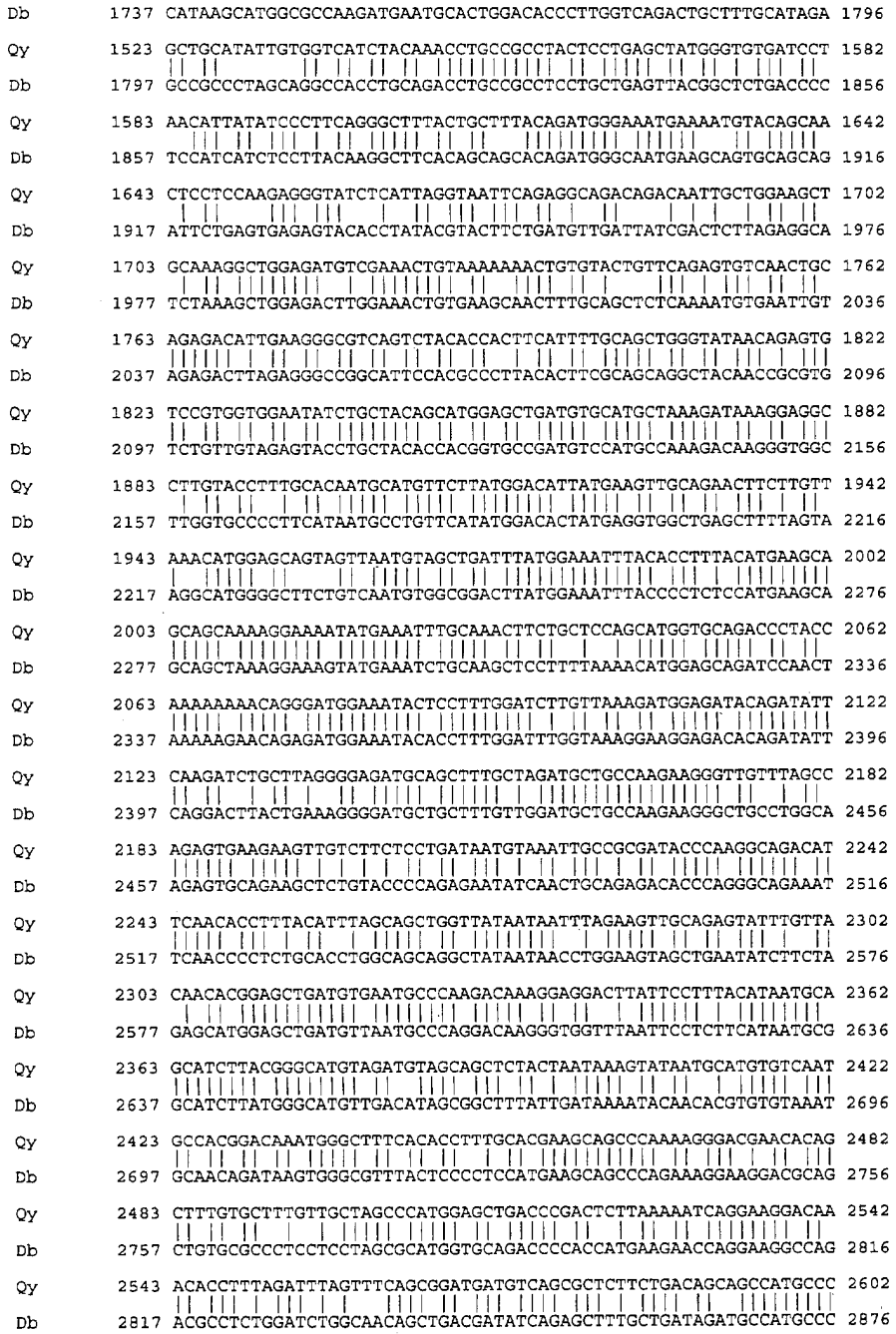
RESULT 38
AF082556
LOCUS AF082556 4134 bp mRNA linear PRI 30-NOV-1998
DEFINITION Homo sapiens TRF1-interacting ankyrin-related ADP-ribose polymerase mRNA, complete cds.
ACCESSION AF082556
VERSION AF082556.1 GI:3929218
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4134)
AUTHORS Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
TITLE Tankyrase, a poly(ADP-ribose) polymerase at human telomeres
JOURNAL Science 282 (5393), 1484-1487 (1998)
MEDLINE 99040105
PUBMED 9822378
REFERENCE 2 (bases 1 to 4134)
AUTHORS Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-1998) Laboratory for Cell Biology and Genetics, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

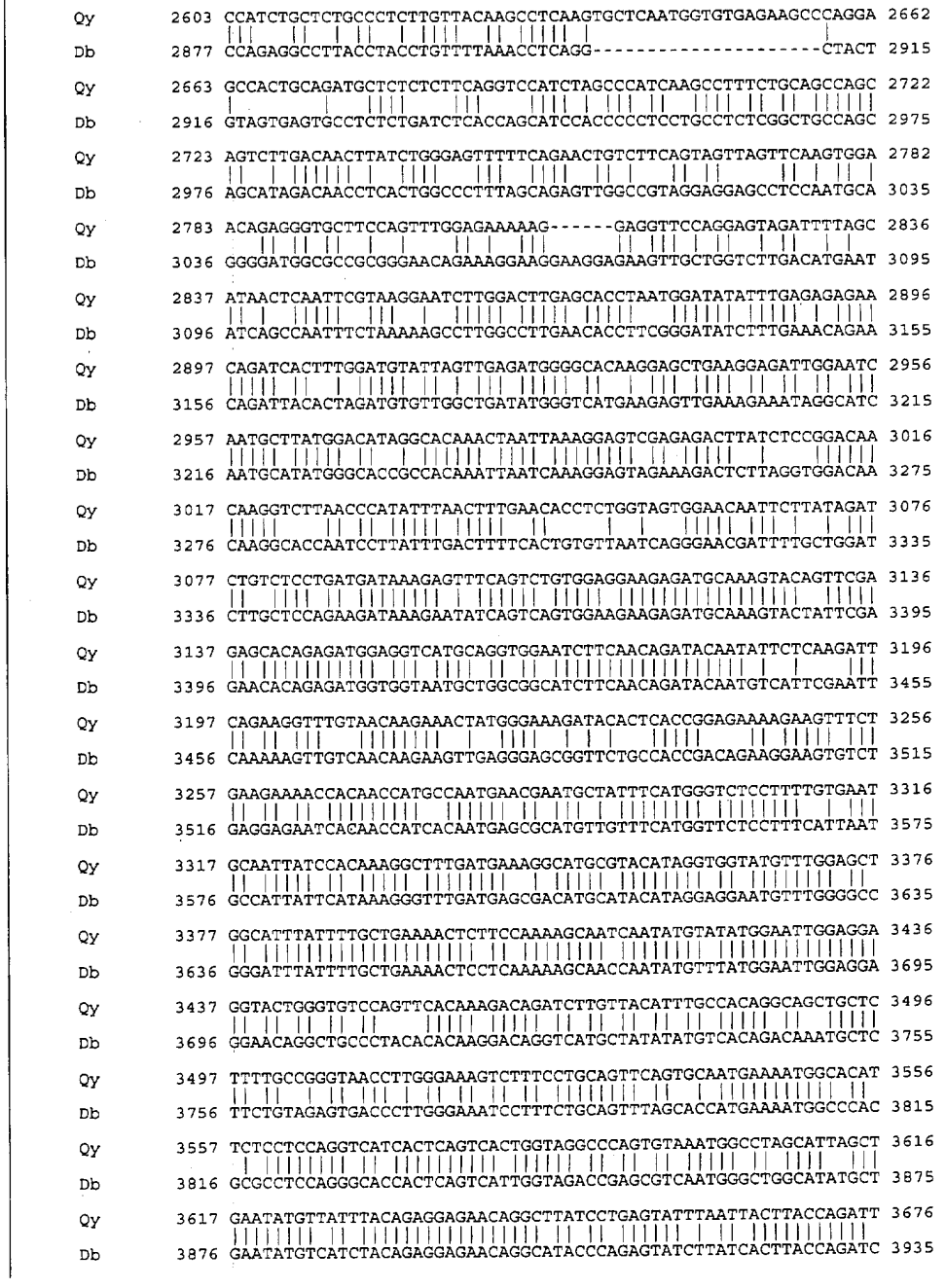
FEATURES Location/Qualifiers
source 1. .4134
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="testis"
CDS 6. .3989
/note="tankyrase"
/codon_start=1
/product="TRF1-interacting ankyrin-related ADP-ribose polymerase"
/protein_id="AAC79841.1"
/db_xref="GI:3929219"
/translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPLSPGLAPGTT
PASPTASGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGTSCCSTTSTICTVAAA
PVVPAVSTSSAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPGSSLAESPEAAGVS
STAPLGPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSS
PLHFAAGFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPN
ARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEY
KKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHG
ADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEV
CSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKT
LALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQG
FTAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEG
RHSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHG
ASVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQ
DLLKGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYL
LEHGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKG
RTQLCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVV
SASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNAGDGAAGTERKEGEVAGLDMN
ISQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLG
GQQGTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYN
VIRIQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYI
GGMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQ
FSTMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTA
TAAEQKT"

ORIGIN

Query Match 47.8%; Score 1813.4; DB 9; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716
Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776
Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836
Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896
Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956
Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016
Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db        1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076
Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db        1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136
Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db        1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196
Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db        1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256
Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db        1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316
Qy        1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db        1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376
Qy        1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db        1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436
Qy        1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db        1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496
Qy        1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db        1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556
Qy        1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db        1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616
Qy        1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db        1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676
Qy        1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db        1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736
Qy        1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db        1737 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1796
```





```
Qy       3677 ATGAGGCCTGAAG 3689
              |||| ||| ||||
Db       3936 ATGAAGCCAGAAG 3948


RESULT 39
AY142108
LOCUS       AY142108               3804 bp    mRNA    linear   VRT 30-MAY-2003
DEFINITION  Gallus gallus tankyrase 1 mRNA, complete cds.
ACCESSION   AY142108
VERSION     AY142108.1  GI:27461954
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 3804)
  AUTHORS   De Rycker,M., Venkatesan,R.N., Wei,C. and Price,C.M.
  TITLE     Vertebrate tankyrase domain structure and sterile alpha motif
            (SAM)-mediated multimerization
  JOURNAL   Biochem. J. 372 (Pt 1), 87-96 (2003)
  MEDLINE   22615216
   PUBMED   12589701
REFERENCE   2  (bases 1 to 3804)
  AUTHORS   Venkatesan,R.N., De Rycker,M. and Price,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-AUG-2002) Department of Molecular Genetics,
            University of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH
            45267-0524, USA
FEATURES             Location/Qualifiers
     source          1. .3804
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9031"
     CDS             1. .3801
                     /codon_start=1
                     /product="tankyrase 1"
                     /protein_id="AAN41651.1"
                     /db_xref="GI:27461955"
                     /translation="MAAPPRRSQHHHHHHGPPPPPGPASPPAAASPPRSPSLAPAELG
                     PAAQRHSLAGPEGEAPPDAERPPAPECSEGAAPGPPPGSGSSSGSSASSSSSSSSTSS
                     SVASSPAAESPEAAGPSGAFRELLEACRNGDVTRVKRLVDAGNVNAKDMAGRKSTPLH
                     FAAGFGRKDVVEHLLQTGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNARD
                     NWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAEAVLTGEYKKD
                     ELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADV
                     HAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSL
                     LLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLAL
                     EIINFKQPQSHETALHCAVAAVHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAEKA
                     HNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLNYGSDPSIISLQGFTA
                     AQIGNEAVQQILSESTPVRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHS
                     TPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASV
                     NVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLL
                     RGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEH
                     GADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQ
                     LCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKLQATVVSAS
                     LISPASTPSCLSAASSIDNLTGPLAELAVGGASNTGDGAAGTERKEGEVSGLDMNITQ
                     FLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQ
                     GTNPYLTFHCVSQGTILLDLAPDDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIR
                     IQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGM
                     FGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFST
                     MKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIVKPEAPSQTGTAA
                     EQKT"
ORIGIN

  Query Match           47.6%;  Score 1808.8;  DB 5;  Length 3804;
  Best Local Similarity  72.4%;  Pred. No. 0;
  Matches 2385;  Conservative    0;  Mismatches  882;  Indels   27;  Gaps    2;

Qy        404 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCA 463
              || |||||  |||||||  || ||  | ||||| ||||  || ||  | || || ||
Db        490 GGGTTTGGAAGGAAAGATGTTGTCGAGCACTTGCTACAGACAGGAGCCAACGTTCATGCT 549

Qy        464 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 523
              ||||||| || || || || || ||||| || || ||||| ||||||||||||||||
Db        550 CGTGATGACGGAGGGCTGATACCCCTTCACAACGCCTGCTCCTTTGGTCATGCTGAAGTG 609

Qy        524 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 583
              || | ||| | || | || |||||||||| |||||| |||| ||||||||| || |||
Db        610 GTTAGTCTGTTACTGTGCCAAGGTGCAGACCCTAATGCTAGAGACAATTGGAACTACACT 669

Qy        584 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 643
              ||| | || || ||||| || || || ||||| ||||| || |||||||| | |||||
Db        670 CCTTTGCACGAGGCTGCTATCAAGGGGAAGATAGATGTGTGTATTGTGCTGCTGCAGCAC 729

Qy        644 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 703
              |||||||| |||| ||| || || || ||||| | ||| ||||| | ||||| ||
Db        730 GGAGCTGATCCAAACATTCGGAACACCGATGGGAAATCTGCGCTGGATCTGGCAGACCCC 789

Qy        704 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 763
              || ||  |||| || || || |||||||| ||||| || |||||| | |||  ||| |||
Db        790 TCAGCAGAAGCTGTACTCACAGGTGAATACAAGAAGGACGAACTCCTGGAAGCTGCTAGG 849

Qy        764 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 823
              ||||| |||||||||||| |||||||| || | || || |||||||  |||||||  |||
Db        850 AGTGGTAATGAAGAAAAACTGATGGCTTTATTGACTCCTTTAAATGTGAACTGCCATGCA 909

Qy        824 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883
              || |||||  | ||||| |||||  | |||||||| || |||||||||    ||
Db        910 AGCGATGGACGCAAGTCCACTCCTTTACACCTAGCAGCAGGCTACAACAGAGTACGAATA 969

Qy        884 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 943
              || || ||  | || || |||| || |||||  || || || || |||||||  ||| ||
Db        970 GTCCAACTTCTGCTTCAGCATGGTGCAGATGTTCACGCAAAGGACAAAGGTGGTCTTGTG 1029

Qy        944 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1003
              ||  | || || || ||||| |||||  || |||||||| || ||  |  |||||
Db       1030 CCTCTTCATAACGCATGTTCATATGGACACTATGAAGTTACAGAGCTACTACTTAAGCAC 1089

Qy       1004 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1063
              ||||| || || |||||||||||| | |||||||| |||||  | |||||||| ||||
Db       1090 GGTGCATGCGTTAATGCAATGGACCTCTGGCAATTTACTCCCTTGCATGAGGCTGCTTCC 1149

Qy       1064 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1123
              |||||| | || ||||| ||||| || |  ||| |||||||| || ||  |  ||
Db       1150 AAGAACCGTGTAGAAGTCTGTTCCCTCTTGCTGAGTCACGGTGCAGATCCCACGTTAGTC 1209

Qy       1124 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1183
              || || ||  || ||||  | ||  |||||||  |||||  ||  |||  || ||  |
Db       1210 AACTGCCATGGAAAGAGTGCGGTCGATATGGCTCCGACACCTGAGCTAAGGGAGAGGCTG 1269

Qy       1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1243
              | |||||||| |||||  || || ||  ||||||||  ||  ||||  ||  |||  |  |
Db       1270 ACCTATGAATTCAAAGGGCATTCTTTACTGCAAGCAGCCAGAGAGGCAGATTTAGCCAAA 1329

Qy       1244 ATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1303
              | ||||||    ||  |||| ||  | |||||  |||||  ||  ||  |||||||  |
Db       1330 GTGAAAAAAAACACTTGCTTTGGAGATCATCAATTTTAAGCAGCCACAGTCACATGAGACT 1389

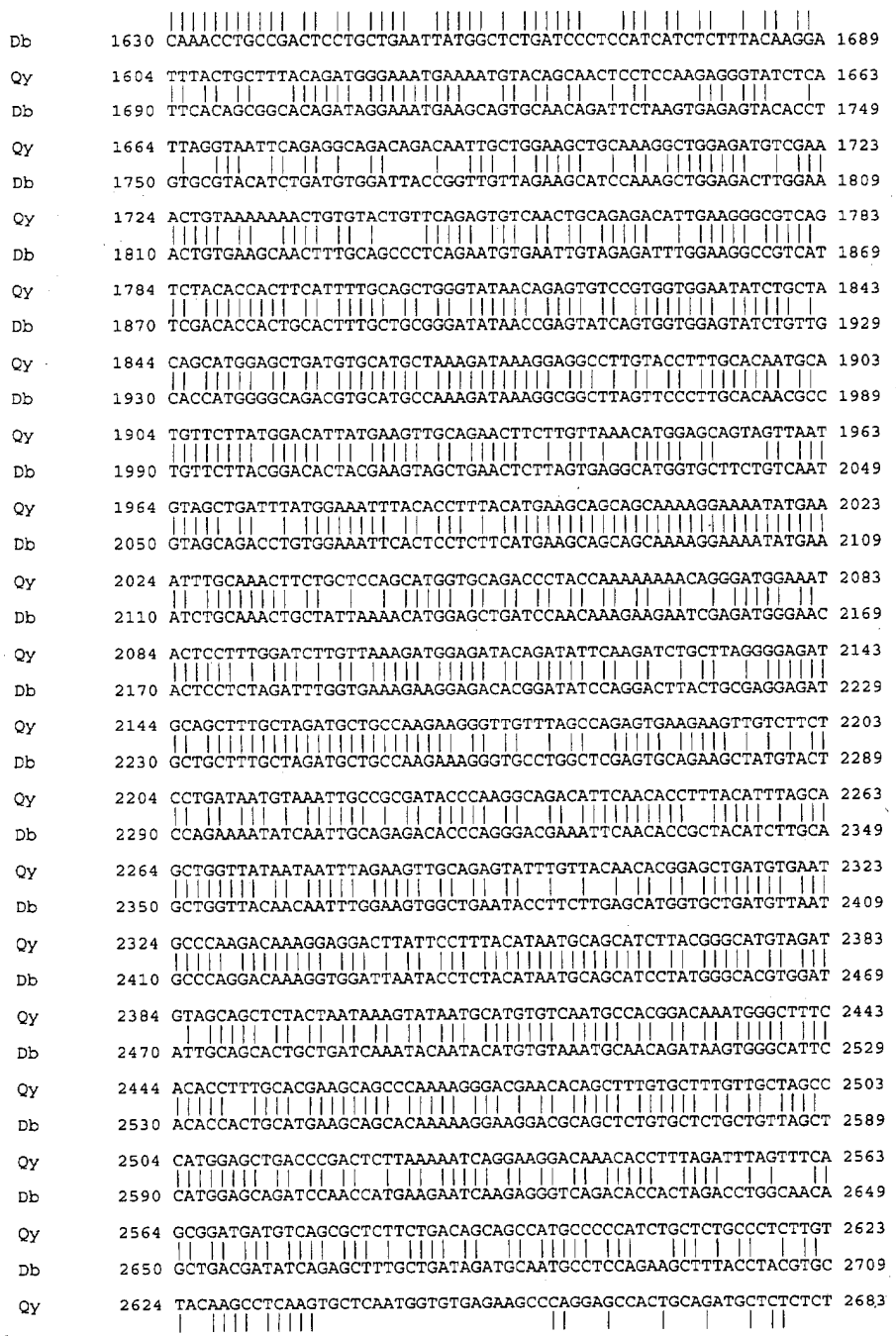
Qy       1304 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1363
              ||  |||| |||||||  ||  ||  |||||||  | ||||||  |  |||||  |||
Db       1390 GCTCTGCACTGTGCTGTGGCCGCTGTGCATCCCAAGCGGAAGCAAGTTACAGAACTATTG 1449

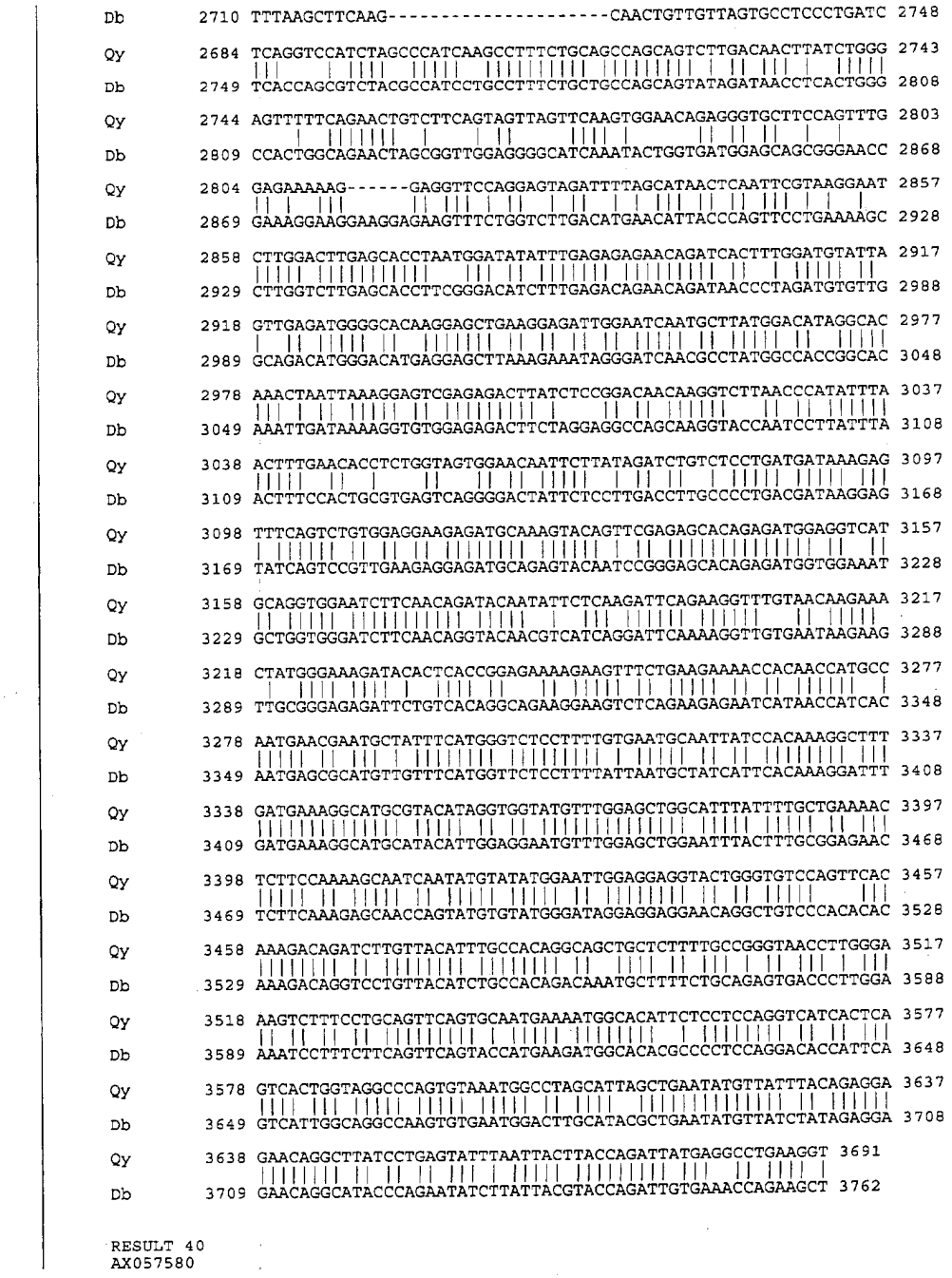
Qy       1364 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1423
              || || |||||||| ||| | |||||  ||  |||||  ||| |||| ||| ||  ||
Db       1450 CTCAGGAAAGGAGCCAACGTTAATGAGAAAAACAAAGATTTCATGACACCTTTGCATGTT 1509

Qy       1424 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1483
              ||  |||||||||||||||||| | |||||  ||  || |||| |||||||  |
Db       1510 GCCGCTGAGAAAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCATGGAGCAAAGATG 1569

Qy       1484 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1543
              || || |||||  |  |||||||| ||  |  | |||||  |||||  |  |||||  |
Db       1570 AACGCACTGGACACACTTGGTCAAACAGCATTGCACAGGGCTGCTTTAGCAGGTCACTTG 1629

Qy       1544 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1603
```





RESULT 40
AX057580

```
LOCUS       AX057580                3254 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 3 from Patent WO0077225.
ACCESSION   AX057580
VERSION     AX057580.1  GI:12310302
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Chi,N.W. and Lodish,H.F.
  TITLE     A novel insulin signaling molecule
  JOURNAL   Patent: WO 0077225-A 3 21-DEC-2000;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; THE GENERAL
            HOSPITAL CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .3254
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"
ORIGIN

  Query Match             47.3%;  Score 1794.8;  DB 6;  Length 3254;
  Best Local Similarity   76.2%;  Pred. No. 0;
  Matches 2518;  Conservative    0;  Mismatches  232;  Indels  554;  Gaps    3;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db        505 AGGTTTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGCAAGC 564

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db        565 GCGTGATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTGAAGT 624

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db        625 AGTCAATCTTCTTTTGCAACACGGTGCAGACCCAAATGCTCGGGATAATTGGAATTATAC 684

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db        685 TCCTCTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTACAGCA 744

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db        745 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGACCC 804

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db        805 ATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGAAAGATGAACTTCTGGAAAGTGCCAG 864

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db        865 GAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCCATGC 924

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db        925 AAGTGATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGATACAACAGAGTAAAAAT 984

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db        985 TGTACAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATCTGGT 1044

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db       1045 ACCACTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCA 1104

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db       1105 CGGTGCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAGCTTC 1164

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db       1165 TAAGAACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCAACCCTGCT 1224

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db       1225 GAACTGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAAGATT 1284

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db       1285 ATCGTATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTACAAG 1344

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db       1345 AATCAAAAAACATCTTTCTCTGGAAATGGTGAATTTCAAACATCCTCAGACACATGAAAC 1404

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db       1405 AGCACTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAGATATGTGAACTGCT 1464

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db       1465 GCTCAGAAAGGGAGCAAACACCAACGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1524

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db       1525 GGCTTCTGAGAACGCTCACAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1584

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db       1585 TAATGCTTTGGATAGTCTTGGACAGACGTCATTACACAGAGCTGCACACTGTGGTCACCT 1644

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db       1645 GCAGACCTGCCGCCTGCTCCTGAGCTATGGGTGTGATCCAAACATCATATCCCTTCAGGG 1704

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db       1705 TTTCACCGCCTTGCAAATGGGAAATGAAAATGTGCAGCAGCTGCTCCAAG---------- 1754

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db       1755 ----------------------------------------------------------GA 1756

Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db       1757 AACTGTAAAAAAACTCTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCA 1816

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db       1817 GTCAACCCCACTCCACTTTGCAGCTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCT 1876

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db       1877 GCAACATGGAGCTGATGTTCATGCTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGC 1936

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       1937 ATGTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAA 1996

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       1997 TGTCGCTGACTTGTGGAAGTTCACACCTTTACATGAAGCTGCAGCAAAAGGAAAATATGA 2056

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db       2057 AATATGCAAACTTCTACTCCAGCATGGTGCAGACCCTACAAAGAAAAACAGAGATGGTAA 2116

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       2117 TACTCCTCTGGATCTTGTTAAAGATGGAG----------------------------GA 2147

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       2148 TGCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTC 2207
```

```
Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2208 TCCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACACCTTTACATTTGGC 2267

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       2268 AGCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAA 2327

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2328 TGCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTAGA 2387

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2388 TGTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTT 2447

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2448 CACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGC 2507

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2508 CCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTTC 2567

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2568 AGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCC---------------------- 2606

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2607 ------------------------------------------------------------ 2606

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2607 ------------------------------------------------------------ 2606

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2607 ------------------------------------------------------------ 2606

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       2607 ------------------------------------------------------------ 2606

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2607 ------------------------------------------------------------ 2606

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2607 ------------------------------------------------------------ 2606

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2607 ------------------------------------------------------------ 2606

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       2607 ----------------GGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCA 2650

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       2651 GTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGG 2710

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       2711 CGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTG 2770

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       2771 GGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCAAACGA 2830

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       2831 AAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTGATGA 2890

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       2891 AAGGCATGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTC 2950

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       2951 CAAAAGCAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTCCAATTCACAAAGA 3010

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3011 CAGATCGTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAAAGTC 3070

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3071 TTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGGTCAC 3130

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3131 TGGCCGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3190

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3191 GGCTTATCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAGGTATGGTTGATGG 3250

Qy       3703 ATAA 3706
Db       3251 CTGA 3254
```

```
RESULT 41
AF082557
LOCUS       AF082557               8901 bp    mRNA    linear   PRI 30-NOV-1998
DEFINITION  Homo sapiens TRF1-interacting ankyrin-related ADP-ribose polymerase
            mRNA, partial cds.
ACCESSION   AF082557
VERSION     AF082557.1  GI:3929220
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8901)
  AUTHORS   Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
  TITLE     Tankyrase, a poly(ADP-ribose) polymerase at human telomeres
  JOURNAL   Science 282 (5393), 1484-1487 (1998)
  MEDLINE   99040105
   PUBMED   9822378
REFERENCE   2  (bases 1 to 8901)
  AUTHORS   Smith,S., Giriat,I., Schmitt,A. and de Lange,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1998) Laboratory for Cell Biology and Genetics,
            The Rockefeller University, 1230 York Avenue, New York, NY 10021,
            USA
FEATURES             Location/Qualifiers
     source          1. .8901
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="HeLa"
     CDS             <1. .3285
                     /note="tankyrase"
                     /codon_start=1
                     /product="TRF1-interacting ankyrin-related ADP-ribose
                     polymerase"
                     /protein_id="AAC79842.1"
                     /db_xref="GI:3929221"
                     /translation="HLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPNA
                     RDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEYK
                     KDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGA
                     DVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVC
```

SLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTL
ALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAAE
RAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGF
TAAQMGNEAVQQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGR
HSTPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGA
SVNVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQD
LLKGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLL
EHGADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGR
TQLCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVVS
ASLISPASTPSCLSAASSIDNLTGPLAELAVGGASNAGDGAAGTERKEGEVAGLDMNI
SQFLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGG
QQGTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNV
IRIQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIG
GMFGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQF
STMKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTAT
AAEQKT"

ORIGIN

Query Match 47.3%; Score 1794.8; DB 9; Length 8901;
Best Local Similarity 72.5%; Pred. No. 0;
Matches 2365; Conservative 0; Mismatches 872; Indels 27; Gaps 2;

```
Qy       432 ATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTC 491
Db         2 ACTTACTACAGATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTC 61

Qy       492 ATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAG 551
Db        62 ATAATGCCTGTTCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTG 121

Qy       552 ACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAA 611
Db       122 ATCCAAATGCCAGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGA 181

Qy       612 AGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAG 671
Db       182 AGATCGATGTGTGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTG 241

Qy       672 ATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAAT 731
Db       242 ATGGGAAATCAGCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAAT 301

Qy       732 ATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTC 791
Db       302 ACAAGAAAGACGAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTT 361

Qy       792 TACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTAC 851
Db       362 TACTGACTCCTCTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTAC 421

Qy       852 ATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTG 911
Db       422 ATCTAGCAGCGGGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTG 481

Qy       912 ATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTC 971
Db       482 ATGTTCATGCAAAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGAC 541

Qy       972 ATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGT 1031
Db       542 ATTATGAAGTCACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCT 601

Qy      1032 GGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTC 1091
Db       602 GGCAGTTTACTCCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGT 661

Qy      1092 TCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACT 1151
Db       662 TACTTAGCCATGGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATA 721

Qy      1152 TGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGC 1211
Db       722 TGGCTCCAACTCCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTAC 781

Qy      1212 TGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGG 1271
Db       782 TACAAGCAGCCAGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCA 841

Qy      1272 TGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCAT 1331
Db       842 TTAATTTCAAACAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGC 901

Qy      1332 ATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAA 1391
Db       902 ATCCCAAACGTAAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAA 961

Qy      1392 AGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTG 1451
Db       962 AAAATAAAGATTTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCA 1021

Qy      1452 TTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTT 1511
Db      1022 TGGAAGTTCTGCATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTG 1081

Qy      1512 CTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATG 1571
Db      1082 CTTTGCATAGAGCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACG 1141

Qy      1572 GGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAA 1631
Db      1142 GCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAG 1201

Qy      1632 ATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAAT 1691
Db      1202 CAGTGCAGCAGATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGAC 1261

Qy      1692 TGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGA 1751
Db      1262 TCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAA 1321

Qy      1752 GTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGT 1811
Db      1322 ATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCT 1381

Qy      1812 ATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAG 1871
Db      1382 ACAACCGCGTGTCTGTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAG 1441

Qy      1872 ATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAG 1931
Db      1442 ACAAGGGTGGCTTGGTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTG 1501

Qy      1932 AACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTT 1991
Db      1502 AGCTTTTAGTAAGGCATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTC 1561

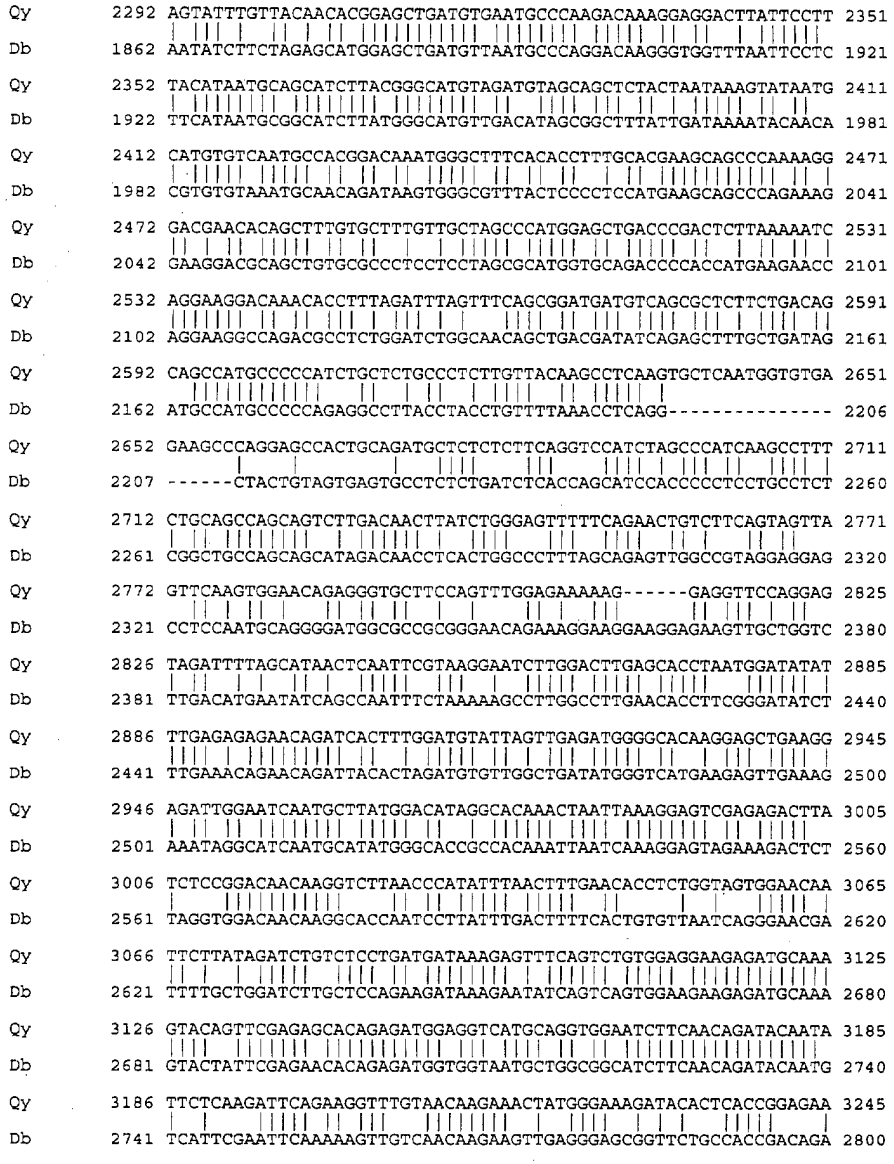
Qy      1992 TACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTG 2051
Db      1562 TCCATGAAGCAGCAGCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAG 1621

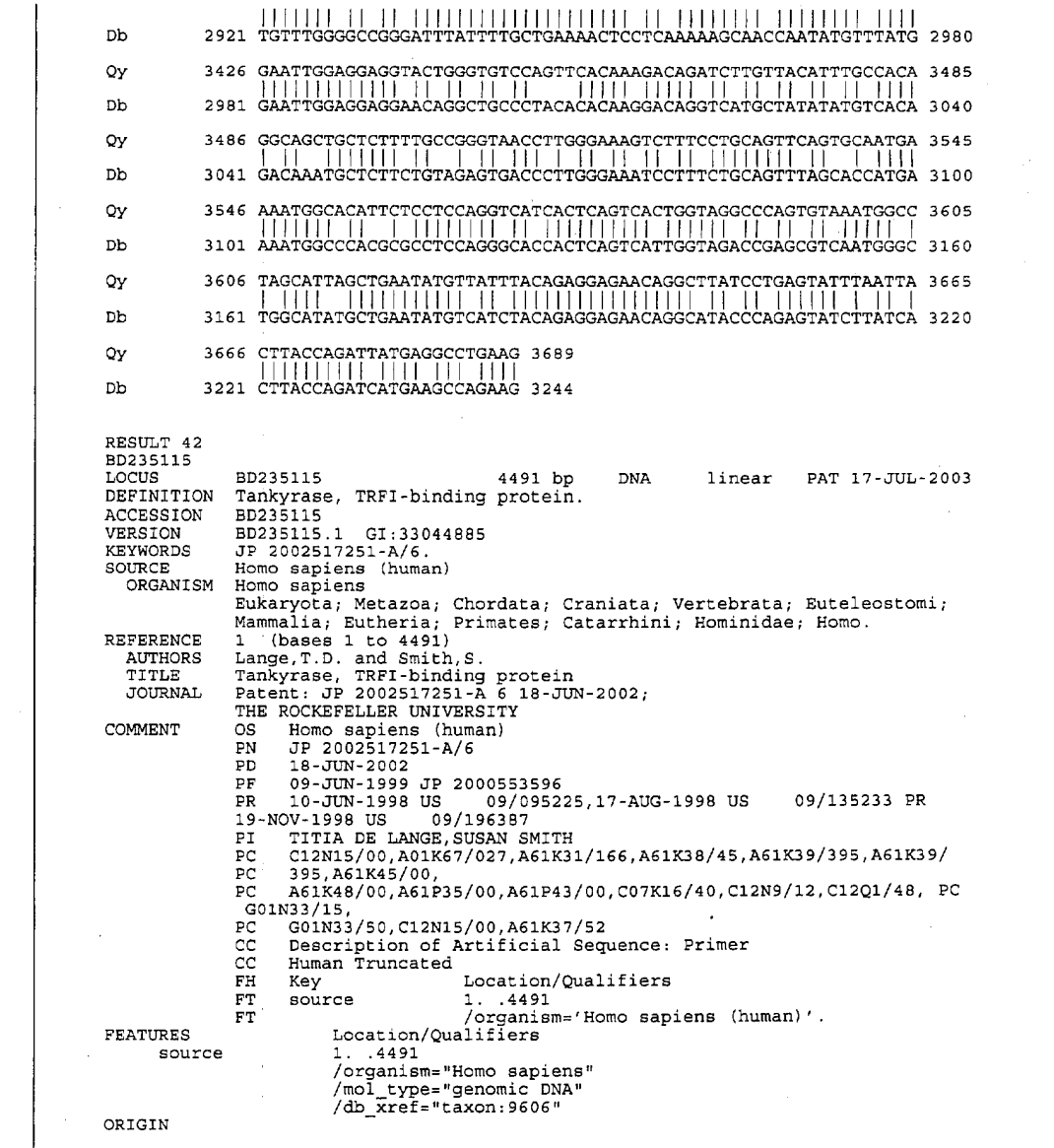
Qy      2052 CAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAG 2111
Db      1622 CAGATCCAACTAAAAAGAACAGAGATGGAAATACACCTTTGGATTTGGTAAAGGAAGGAG 1681

Qy      2112 ATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGG 2171
Db      1682 ACACAGATATTCAGGACTTACTGAAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGG 1741

Qy      2172 GTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCC 2231
Db      1742 GCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCC 1801

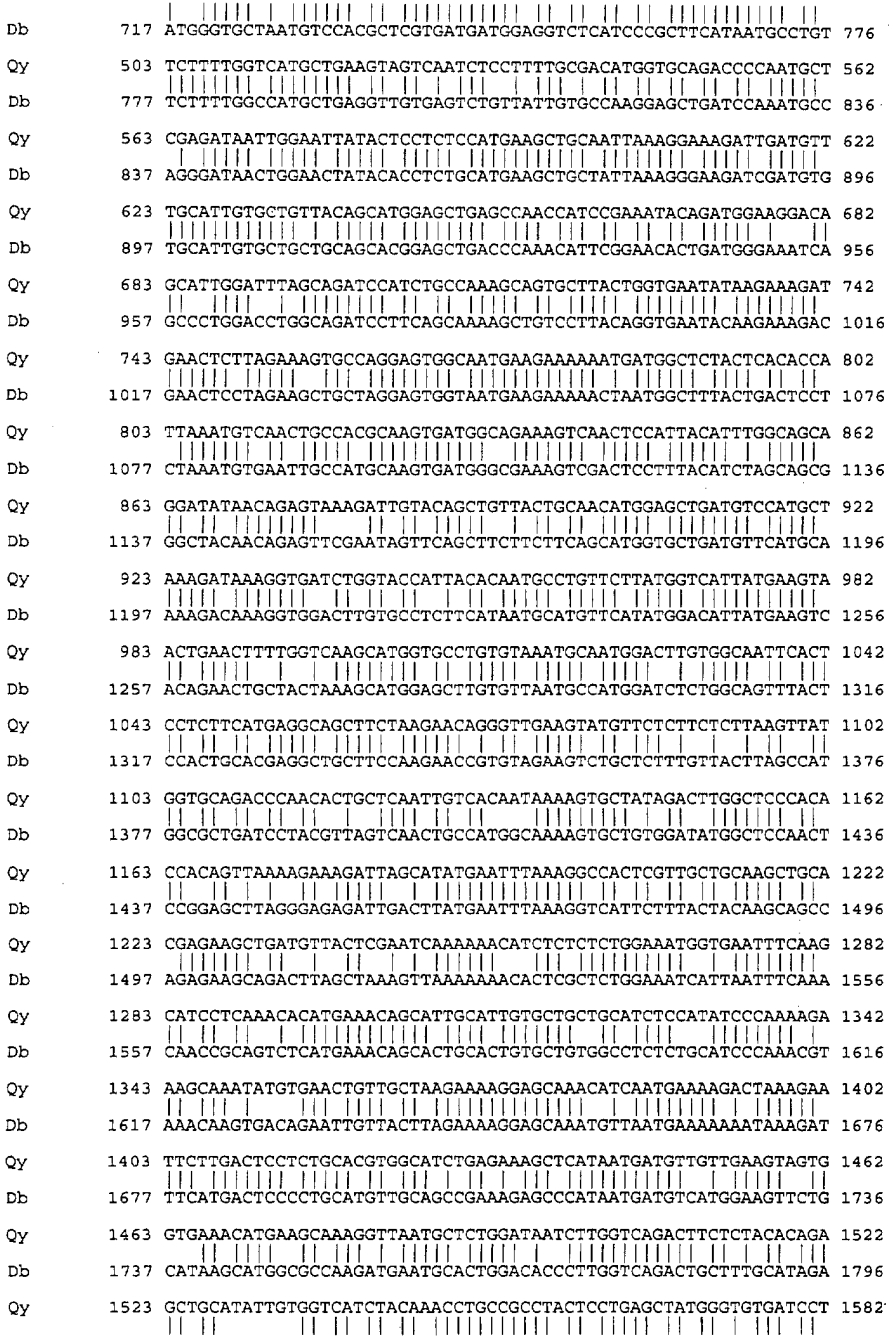
Qy      2232 AAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAG 2291
Db      1802 AGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTG 1861
```

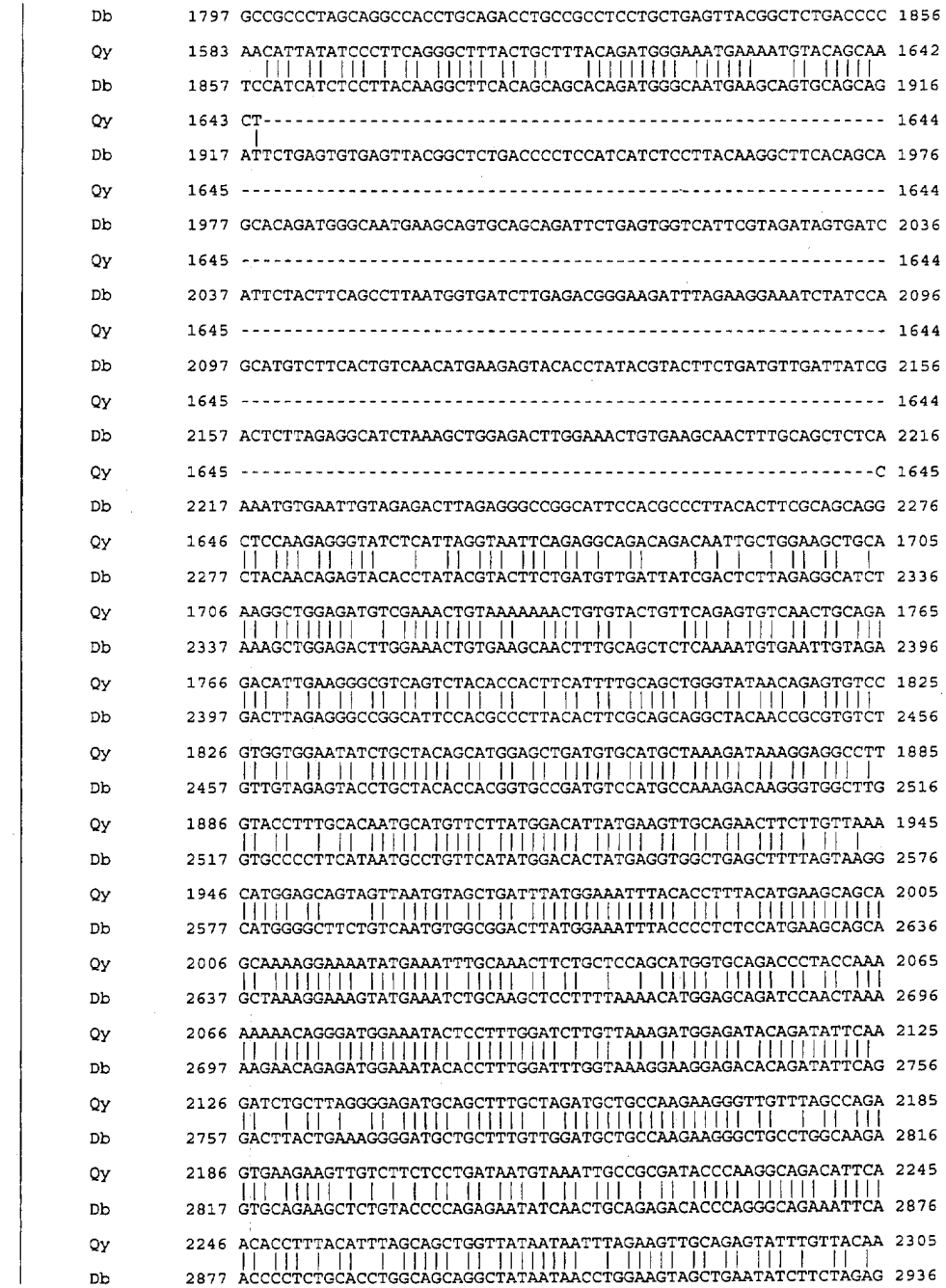


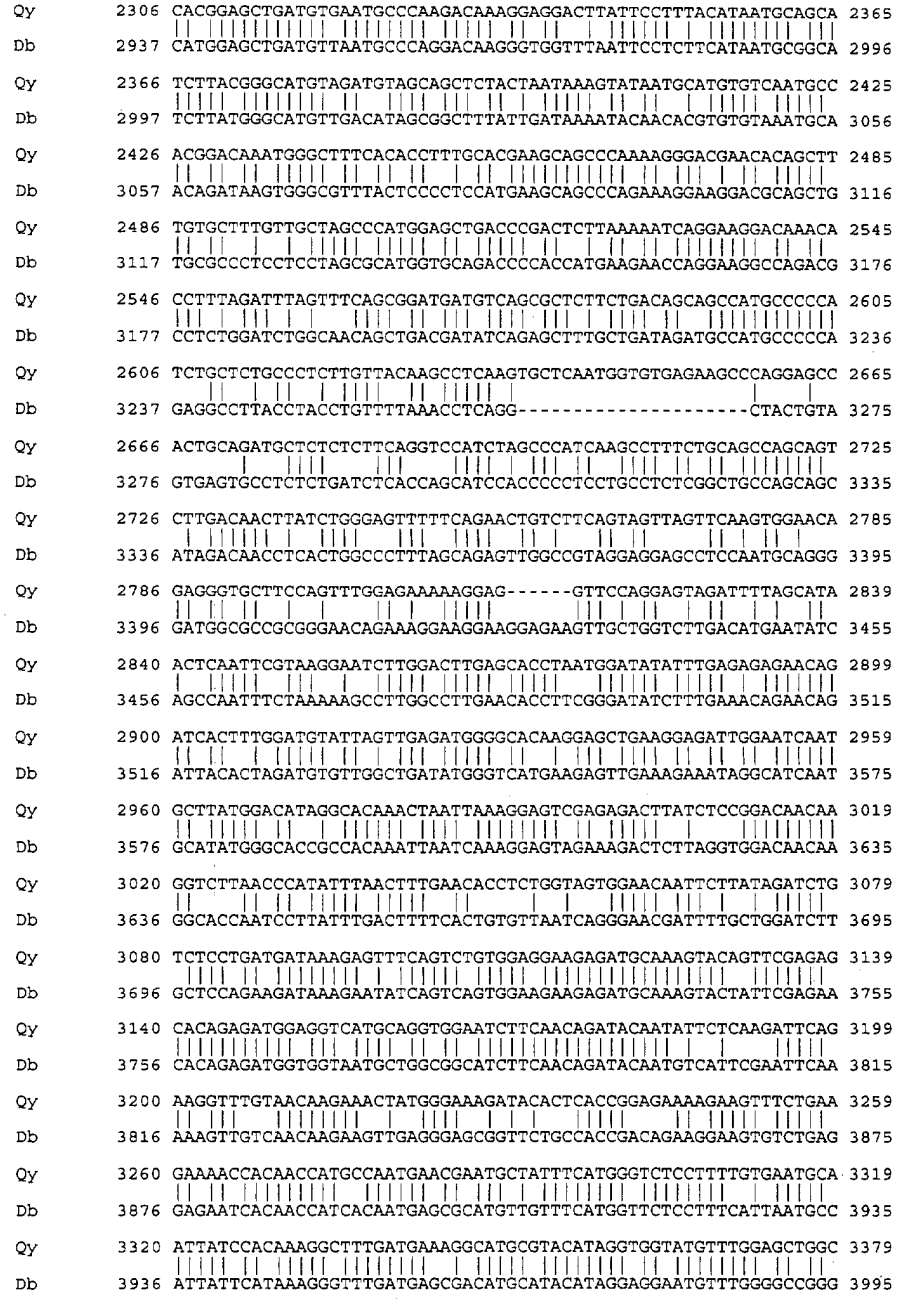


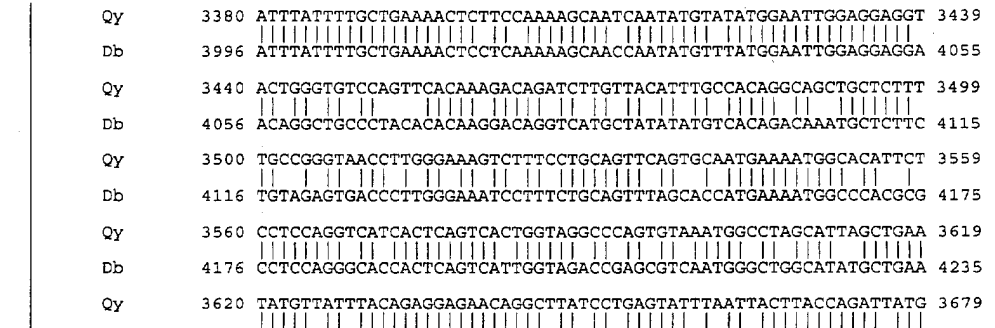
RESULT 42
BD235115
LOCUS BD235115 4491 bp DNA linear PAT 17-JUL-2003
DEFINITION Tankyrase, TRFI-binding protein.
ACCESSION BD235115
VERSION BD235115.1 GI:33044885
KEYWORDS JP 2002517251-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4491)
AUTHORS Lange,T.D. and Smith,S.
TITLE Tankyrase, TRFI-binding protein
JOURNAL Patent: JP 2002517251-A 6 18-JUN-2002;
THE ROCKEFELLER UNIVERSITY
COMMENT OS Homo sapiens (human)
PN JP 2002517251-A/6
PD 18-JUN-2002
PF 09-JUN-1999 JP 2000553596
PR 10-JUN-1998 US 09/095225,17-AUG-1998 US 09/135233 PR
19-NOV-1998 US 09/196387
PI TITIA DE LANGE,SUSAN SMITH
PC C12N15/00,A01K67/027,A61K31/166,A61K38/45,A61K39/395,A61K39/
PC 395,A61K45/00,
PC A61K48/00,A61P35/00,A61P43/00,C07K16/40,C12N9/12,C12Q1/48, PC
G01N33/15,
PC G01N33/50,C12N15/00,A61K37/52
CC Description of Artificial Sequence: Primer
CC Human Truncated
FH Key Location/Qualifiers
FT source 1. .4491
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .4491
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 38.2%; Score 1449.6; DB 6; Length 4491;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;


Db 717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776
Qy 503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db 777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836
Qy 563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db 837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896
Qy 623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db 897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956
Qy 683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db 957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016
Qy 743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db 1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076
Qy 803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db 1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136
Qy 863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db 1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196
Qy 923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db 1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256
Qy 983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db 1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316
Qy 1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db 1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376
Qy 1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db 1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436
Qy 1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db 1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496
Qy 1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db 1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556
Qy 1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db 1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616
Qy 1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db 1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676
Qy 1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db 1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736
Qy 1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db 1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796
Qy 1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582

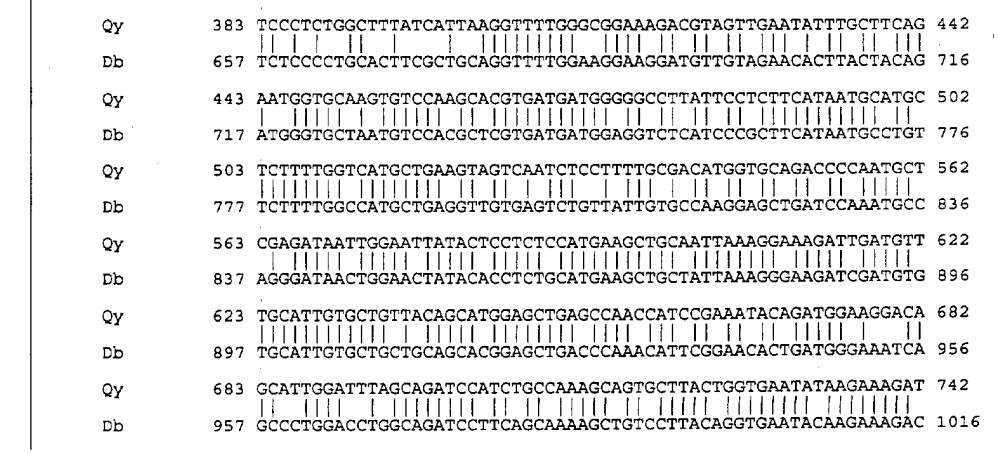

Db 1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856
Qy 1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
Db 1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916
Qy 1643 CT---------- 1644
Db 1917 ATTCTGAGTGTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1976
Qy 1645 ---------- 1644
Db 1977 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGGTCATTCGTAGATAGTGATC 2036
Qy 1645 ---------- 1644
Db 2037 ATTCTACTTCAGCCTTAATGGTGATCTTGAGACGGGAAGATTTAGAAGGAAATCTATCCA 2096
Qy 1645 ---------- 1644
Db 2097 GCATGTCTTCACTGTCAACATGAAGAGTACACCTATACGTACTTCTGATGTTGATTATCG 2156
Qy 1645 ---------- 1644
Db 2157 ACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCA 2216
Qy 1645 ----------C 1645
Db 2217 AAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGG 2276
Qy 1646 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1705
Db 2277 CTACAACAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCATCT 2336
Qy 1706 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1765
Db 2337 AAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGTAGA 2396
Qy 1766 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1825
Db 2397 GACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTGTCT 2456
Qy 1826 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1885
Db 2457 GTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAGACAAGGGTGGCTTG 2516
Qy 1886 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1945
Db 2517 GTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTGAGCTTTTAGTAAGG 2576
Qy 1946 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2005
Db 2577 CATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTCTCCATGAAGCAGCA 2636
Qy 2006 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2065
Db 2637 GCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAGCAGATCCAACTAAA 2696
Qy 2066 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2125
Db 2697 AAGAACAGAGATGGAAATACACCTTTGGATTTGGTAAAGGAAGGAGACACAGATATTCAG 2756
Qy 2126 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2185
Db 2757 GACTTACTGAAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGGGCTGCCTGGCAAGA 2816
Qy 2186 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2245
Db 2817 GTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAATTCA 2876
Qy 2246 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2305
Db 2877 ACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTTCTAGAG 2936

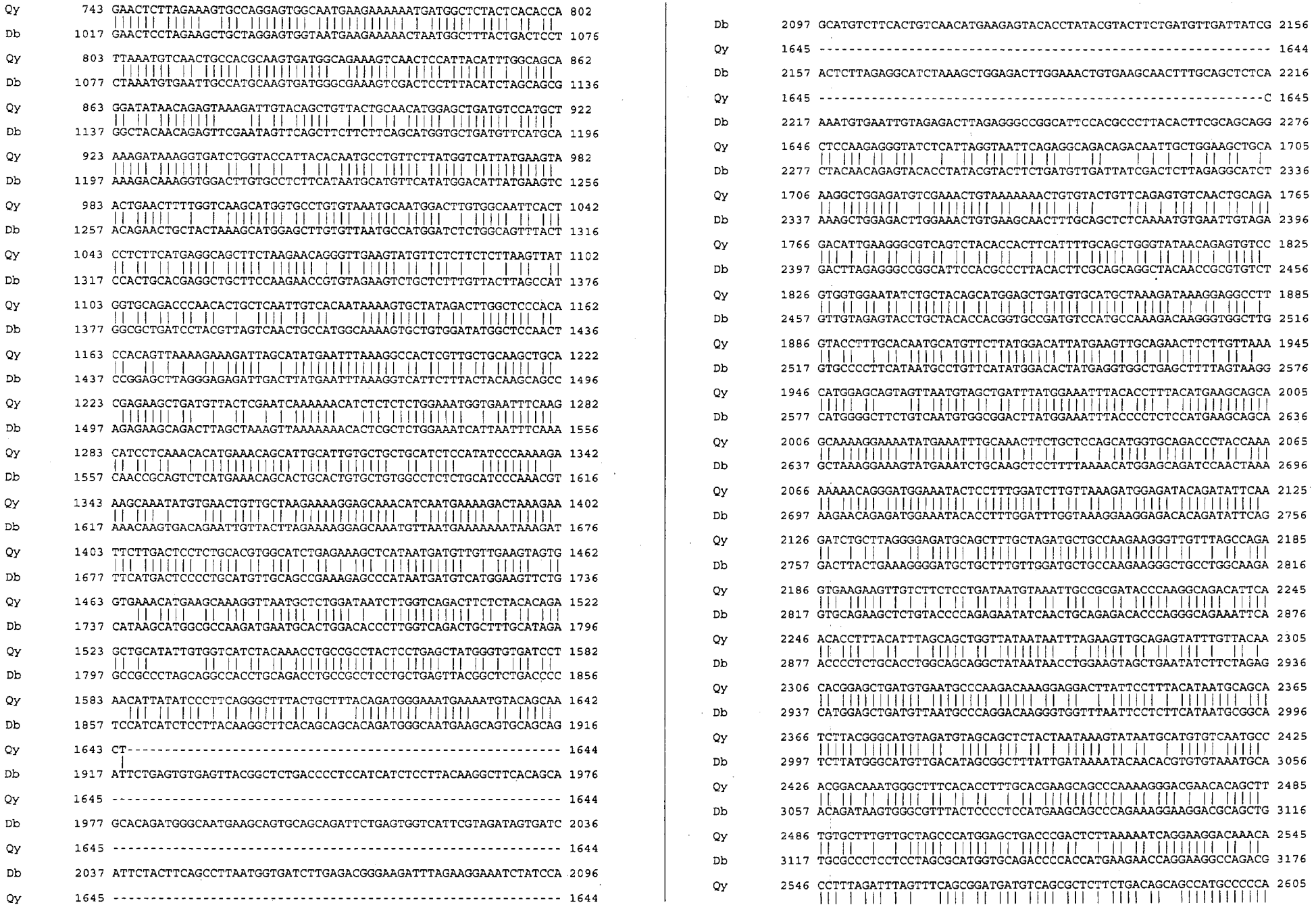




RESULT 43
AR274526
LOCUS AR274526 4491 bp mRNA linear PAT 10-APR-2003
DEFINITION Sequence 7 from patent US 6506587.
ACCESSION AR274526
VERSION AR274526.1 GI:29707026
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 4491)
AUTHORS De Lange,T. and Smith,S.
TITLE TRF 1 binding protein, methods of use thereof
JOURNAL Patent: US 6506587-A 7 14-JAN-2003;
FEATURES Location/Qualifiers
source 1. .4491
/organism="unknown"
/mol_type="mRNA"
ORIGIN

Query Match 38.2%; Score 1449.6; DB 6; Length 4491;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 2397; Conservative 0; Mismatches 889; Indels 384; Gaps 3;





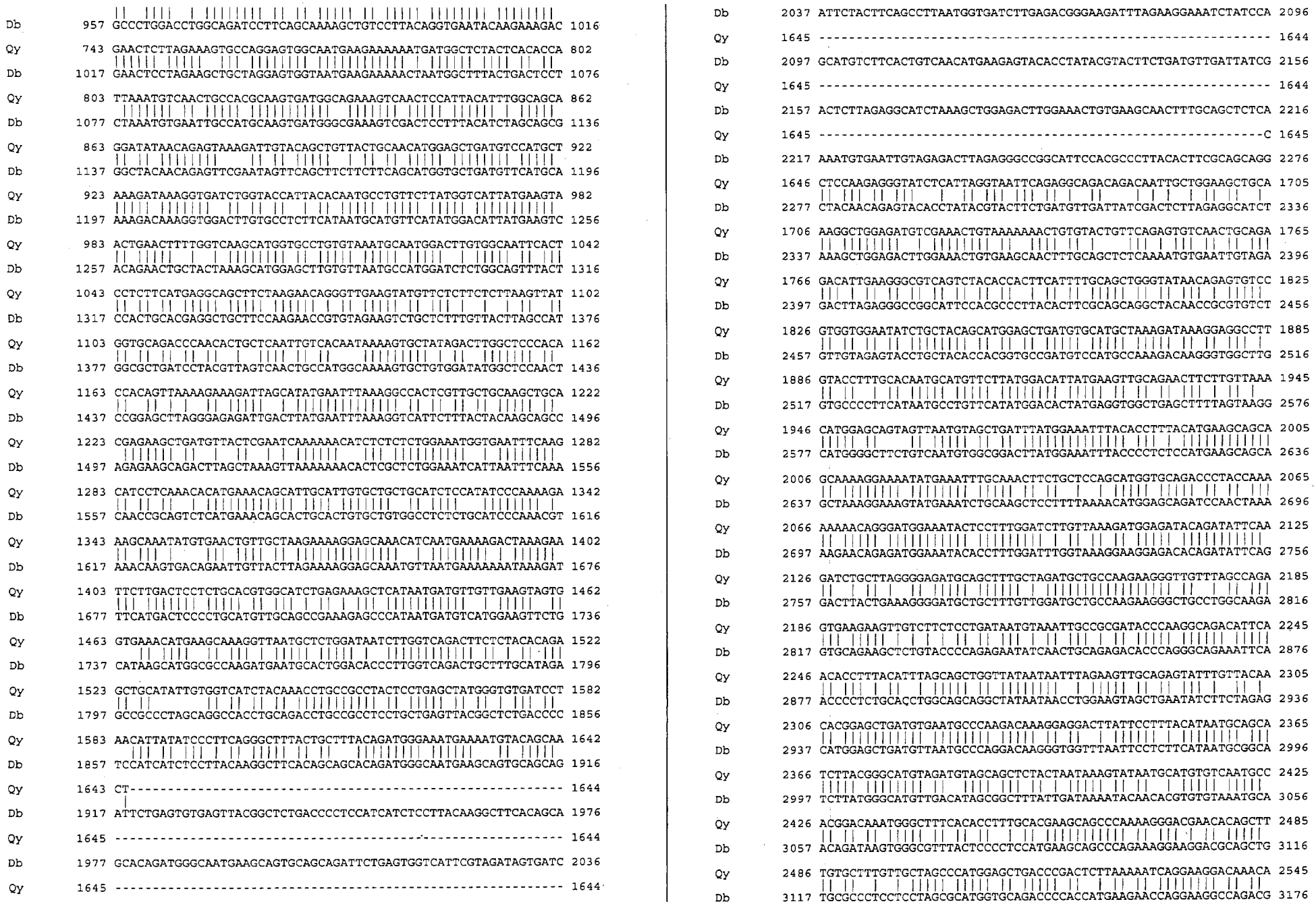
```
Db       3177 CCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCCCCCA 3236
Qy       2606 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2665
Db       3237 GAGGCCTTACCTACCTGTTTTAAACCTCAGG--------------------CTACTGTA 3275
Qy       2666 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2725
Db       3276 GTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGCAGC 3335
Qy       2726 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2785
Db       3336 ATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCAGGG 3395
Qy       2786 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAG------GTTCCAGGAGTAGATTTTAGCATA 2839
Db       3396 GATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAATATC 3455
Qy       2840 ACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAG 2899
Db       3456 AGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAACAG 3515
Qy       2900 ATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAAT 2959
Db       3516 ATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATCAAT 3575
Qy       2960 GCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAA 3019
Db       3576 GCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAA 3635
Qy       3020 GGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTG 3079
Db       3636 GGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTT 3695
Qy       3080 TCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAG 3139
Db       3696 GCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAA 3755
Qy       3140 CACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAG 3199
Db       3756 CACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAA 3815
Qy       3200 AAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAA 3259
Db       3816 AAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAG 3875
Qy       3260 GAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCA 3319
Db       3876 GAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCC 3935
Qy       3320 ATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGC 3379
Db       3936 ATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGG 3995
Qy       3380 ATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGT 3439
Db       3996 ATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGA 4055
Qy       3440 ACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTT 3499
Db       4056 ACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTC 4115
Qy       3500 TGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCT 3559
Db       4116 TGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCG 4175
Qy       3560 CCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAA 3619
Db       4176 CCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAA 4235
Qy       3620 TATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATG 3679
Db       4236 TATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATG 4295
```

```
Qy       3680 AGGCCTGAAG 3689
Db       4296 AAGCCAGAAG 4305
```

```
RESULT 44
AX467067
LOCUS       AX467067               4491 bp    DNA     linear   PAT 16-JUL-2002
DEFINITION  Sequence 7 from Patent WO0229017.
ACCESSION   AX467067
VERSION     AX467067.1  GI:21900395
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   de Lange,T. and Smith,S.
  TITLE     A novel trf1 binding protein, methods of use thereof
  JOURNAL   Patent: WO 0229017-A 7 11-APR-2002;
            THE ROCKEFELLER UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1. .4491
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="cDNA"
     CDS             6. .2027
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD42445.1"
                     /db_xref="GI:21900396"
                     /db_xref="REMTREMBL:CAD42445"
                     /translation="MAASRRSQHHHHHHQQQLQPAPGASAPPPPPPPPLSPGLAPGTT
                     PASPTASGLAPFASPRHGLALPEGDGSRDPPDRPRSPDPVDGTSCCSTTSTICTVAAA
                     PVVPAVSTSSAAGVAPNPAGSGSNNSPSSSSSPTSSSSSSPSSPGSSLAESPEAAGVS
                     STAPLGPGAAGPGTGVPAVSGALRELLEACRNGDVSRVKRLVDAANVNAKDMAGRKSS
                     PLHFAAGFGRKDVVEHLLQMGANVHARDDGGLIPLHNACSFGHAEVVSLLLCQGADPN
                     ARDNWNYTPLHEAAIKGKIDVCIVLLQHGADPNIRNTDGKSALDLADPSAKAVLTGEY
                     KKDELLEAARSGNEEKLMALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHG
                     ADVHAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEV
                     CSLLLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKT
                     LALEIINFKQPQSHETALHCAVASLHPKRKQVTELLLRKGANVNEKNKDFMTPLHVAA
                     ERAHNDVMEVLHKHGAKMNALDTLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQG
                     FTAAQMGNEAVQQILSVSYGSDPSIISLQGFTAAQMGNEAVQQILSGHS"
ORIGIN

  Query Match           38.2%;  Score 1449.6;  DB 6;  Length 4491;
  Best Local Similarity  65.3%;  Pred. No. 0;
  Matches 2397;  Conservative    0;  Mismatches  889;  Indels  384;  Gaps    3;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716
Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776
Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836
Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896
Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956
Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
```



```
Qy       2546 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2605
Db       3177 CCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCCCCCA 3236

Qy       2606 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2665
Db       3237 GAGGCCTTACCTACCTGTTTTAAACCTCAGG----------------------CTACTGTA 3275

Qy       2666 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2725
Db       3276 GTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGCAGC 3335

Qy       2726 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2785
Db       3336 ATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCAGGG 3395

Qy       2786 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAG------GTTCCAGGAGTAGATTTTAGCATA 2839
Db       3396 GATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAATATC 3455

Qy       2840 ACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAG 2899
Db       3456 AGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAACAG 3515

Qy       2900 ATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAAT 2959
Db       3516 ATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATCAAT 3575

Qy       2960 GCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAA 3019
Db       3576 GCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAA 3635

Qy       3020 GGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTG 3079
Db       3636 GGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTT 3695

Qy       3080 TCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAG 3139
Db       3696 GCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAA 3755

Qy       3140 CACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAG 3199
Db       3756 CACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAA 3815

Qy       3200 AAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAA 3259
Db       3816 AAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAG 3875

Qy       3260 GAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCA 3319
Db       3876 GAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCC 3935

Qy       3320 ATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGC 3379
Db       3936 ATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGG 3995

Qy       3380 ATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGT 3439
Db       3996 ATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGA 4055

Qy       3440 ACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTT 3499
Db       4056 ACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTC 4115

Qy       3500 TGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCT 3559
Db       4116 TGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCG 4175

Qy       3560 CCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAA 3619
Db       4176 CCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAA 4235
```

```
Qy       3620 TATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATG 3679
Db       4236 TATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATG 4295

Qy       3680 AGGCCTGAAG 3689
Db       4296 AAGCCAGAAG 4305
```

```
RESULT 45
AX062239
LOCUS       AX062239               1568 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 98 from Patent WO0100849.
ACCESSION   AX062239
VERSION     AX062239.1  GI:12540140
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Christenson,E., Demaggio,A.J., Goldman,P.S. and Mcelligott,D.L.
  TITLE     Tankyrase2 materials and methods
  JOURNAL   Patent: WO 0100849-A 98 04-JAN-2001;
            ICOS CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .1568
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <2. .>1567
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC25059.1"
                     /db_xref="GI:12540141"
                     /db_xref="REMTREMBL:CAC25059"
                     /translation="ARIMSGRRCAGGGAACASAAAEAVEPAARELFEACRNGDVERVK
                     RLVTPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNAC
                     SFGHAEVVNLLLRHGADPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDG
                     RTALDLADPSAKAVLTGEYKKDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPL
                     HLAAGYNRVKIVQLLLQHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAM
                     DLWQFTPLHEAASKNRVEVCSLLLSYGADPTLLNCHNKSAIDLAPTPQLKERLAYEFK
                     GHSLLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRK
                     GANINEKTKEFLTPLHVASEKAHNDVVEVVVKHEAKVNALDNLGQTSLHRAAYCGHLQ
                     TCRLLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGISLGNSEADRQLLEAAKAGDV
                     ETVKKLCTVQSVNC"
ORIGIN

  Query Match           35.8%;  Score 1359.4;  DB 6;  Length 1568;
  Best Local Similarity  99.9%;  Pred. No. 1.6e-306;
  Matches 1360;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db        448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
```

```
                                                
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 19:22:31 ; Search time 9061.83 Seconds
(without alignments)
12512.564 Million cell updates/sec

Title: US-09-843-159B-1
Perfect score: 3797
Sequence: 1 ctttgaagacactggatttc.....cctttgctgaaaaaaaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1609.8	42.4	3417	11	AK048860	AK048860 Mus muscu
	2	1369.8	36.1	1967	11	AK047094	AK047094 Mus muscu
	3	1360.4	35.8	3194	11	AK080913	AK080913 Mus muscu
	4	784.2	20.7	838	9	AU140145	AU140145 AU140145
	5	687.8	18.1	955	14	CB183123	CB183123 AGENCOURT
	6	672.4	17.7	675	14	CB128566	CB128566 K-EST0178
	7	672.2	17.7	1128	12	BM457025	BM457025 AGENCOURT
c	8	665.2	17.5	754	14	CB170279	CB170279 CEY603161
c	9	652	17.2	923	14	CF407574	CF407574 CH3#049_C
	10	649.2	17.1	659	14	R64714	R64714 EST54a09 WA
	11	643.8	17.0	691	12	BG391376	BG391376 602417447
	12	642.8	16.9	759	9	AL601027	AL601027 DKFZp313G
	13	634.4	16.7	950	13	BQ885764	BQ885764 AGENCOURT
	14	632.6	16.7	743	14	CF539235	CF539235 UI-M-GH0-
	15	622.8	16.4	638	13	BX474686	BX474686 DKFZp686J
	16	616.4	16.2	724	14	CF533694	CF533694 UI-M-FY0-
	17	614.2	16.2	747	14	CF727485	CF727485 UI-M-HB0-
	18	605	15.9	648	12	BG875327	BG875327 RC1-BT062
	19	596	15.7	703	14	CD803519	CD803519 UI-M-GV0-
c	20	593.2	15.6	622	10	BE220005	BE220005 hv65h03.x
c	21	584	15.4	616	10	BE467629	BE467629 hz67b11.x
	22	583.8	15.4	793	13	BU466115	BU466115 603371794
	23	578.8	15.2	689	14	CF739820	CF739820 UI-M-HD0-
	24	574	15.1	796	13	BU427328	BU427328 603234467
	25	572.8	15.1	742	14	CF735419	CF735419 UI-M-HB0-
	26	572.6	15.1	913	14	CB201923	CB201923 AGENCOURT
	27	570.4	15.0	956	10	BE910901	BE910901 601661978
	28	568	15.0	702	10	BB339554	BB339554 BB339554
	29	564.2	14.9	725	14	CF538288	CF538288 UI-M-GH0-
c	30	562.8	14.8	585	10	BE222311	BE222311 hu10d01.x
	31	561.2	14.8	681	14	CF728039	CF728039 UI-M-HB0-
	32	557.8	14.7	669	14	CF745978	CF745978 UI-M-GV0-
c	33	557.4	14.7	595	10	BE504175	BE504175 hv86c11.x
	34	552	14.5	640	14	CF530741	CF530741 UI-M-FY0-
	35	549	14.5	553	13	BQ267594	BQ267594 ij90c10.y
	36	545	14.4	691	14	CF729155	CF729155 UI-M-HD0-
	37	544	14.3	1099	12	BG245969	BG245969 602358786
c	38	538.6	14.2	562	10	BE504478	BE504478 hz57a02.x
	39	537.8	14.2	621	14	CF538447	CF538447 UI-M-GI0-
	40	537.6	14.2	703	14	CF734609	CF734609 UI-M-HB0-
	41	535.8	14.1	765	13	BU570530	BU570530 AGENCOURT
	42	535.6	14.1	679	14	CF726657	CF726657 UI-M-HB0-
	43	534.4	14.1	996	10	BF585334	BF585334 602102008
	44	533.2	14.0	769	14	CA317519	CA317519 UI-M-FW0-
	45	528.2	13.9	655	14	CD350434	CD350434 UI-M-FY0-

ALIGNMENTS

RESULT 1
AK048860
LOCUS AK048860 3417 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched library, clone:C230076L23 product:TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE homolog [Homo sapiens], full insert sequence.
ACCESSION AK048860
VERSION AK048860.1 GI:26339607
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374

PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3417)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES Location/Qualifiers
source 1. .3417
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C230076L23"
/db_xref="MGI:2415999"
/db_xref="taxon:10090"
/clone="C230076L23"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
CDS 423. .3353
/note="unnamed protein product; TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE homolog [Homo sapiens] (SPTR|O95271, evidence: FASTY, 95.9%ID, 80%length, match=3167)
putative"
/codon_start=1
/protein_id="BAC33475.1"
/db_xref="GI:26339608"
/translation="MALLTPLNVNCHASDGRKSTPLHLAAGYNRVRIVQLLLQHGADV
HAKDKGGLVPLHNACSYGHYEVTELLLKHGACVNAMDLWQFTPLHEAASKNRVEVCSL
LLSHGADPTLVNCHGKSAVDMAPTPELRERLTYEFKGHSLLQAAREADLAKVKKTLAL
EIINFKQPQSHETALHCAVASLHPKRKQVAELLLRKGANVNEKNKDFMTPLHVAAERA
HNDVMEVLHKHGAKMNALDSLGQTALHRAALAGHLQTCRLLLSYGSDPSIISLQGFTA
AQMGNEAVQQILSESTPMRTSDVDYRLLEASKAGDLETVKQLCSPQNVNCRDLEGRHS
TPLHFAAGYNRVSVVEYLLHHGADVHAKDKGGLVPLHNACSYGHYEVAELLVRHGASV
NVADLWKFTPLHEAAAKGKYEICKLLLKHGADPTKKNRDGNTPLDLVKEGDTDIQDLL
RGDAALLDAAKKGCLARVQKLCTPENINCRDTQGRNSTPLHLAAGYNNLEVAEYLLEH
GADVNAQDKGGLIPLHNAASYGHVDIAALLIKYNTCVNATDKWAFTPLHEAAQKGRTQ
LCALLLAHGADPTMKNQEGQTPLDLATADDIRALLIDAMPPEALPTCFKPQATVVSAS
VISPASTPSCLSAASSIDNLTGPLTDLAVGGASNAGDGAAGAERKEGEVAGLDMNIIQ
FLKSLGLEHLRDIFETEQITLDVLADMGHEELKEIGINAYGHRHKLIKGVERLLGGQQ
GTNPYLTFHCVNQGTILLDLAPEDKEYQSVEEEMQSTIREHRDGGNAGGIFNRYNVIR
IQKVVNKKLRERFCHRQKEVSEENHNHHNERMLFHGSPFINAIIHKGFDERHAYIGGM
FGAGIYFAENSSKSNQYVYGIGGGTGCPTHKDRSCYICHRQMLFCRVTLGKSFLQFST
MKMAHAPPGHHSVIGRPSVNGLAYAEYVIYRGEQAYPEYLITYQIMKPEAPSQTATAA
EQKT"

ORIGIN

Query Match 42.4%; Score 1609.8; DB 11; Length 3417;
Best Local Similarity 72.4%; Pred. No. 0;
Matches 2126; Conservative 0; Mismatches 782; Indels 27; Gaps 2;

```
Qy       761 AGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCAC 820
            ||||||| |||||||||||| | |||||| |||| || || |||||||  || |||||
Db       399 AGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCAT 458

Qy       821 GCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAG 880
            |||||||||||  ||||||| |||||  |||||  |||||||||| || ||||||||   |
Db       459 GCAAGTGATGGACGAAAGTCTACTCCGTTACACCTGGCAGCAGGCTACAACAGAGTTCGG 518

Qy       881 ATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTG 940
            || || |||||| | || || ||||| |||||||||| || |||||  |||||||   ||
Db       519 ATAGTTCAGCTGCTTCTACAGCATGGTGCTGATGTCCACGCGAAAGACAAAGGTGGACTT 578

Qy       941 GTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAG 1000
            |||||  | || ||||| || || || || || ||||| || || |||||  || | |||
Db       579 GTACCTCTTCATAATGCATGCTCATACGGACACTATGAGGTCACAGAACTGCTGCTAAAG 638

Qy      1001 CATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCT 1060
            ||||| || ||||| || || |||||  ||||||| ||||| || || ||||| || |||
Db       639 CATGGAGCTTGTGTCAACGCTATGGATCTGTGGCAGTTCACACCGCTGCATGAAGCTGCT 698

Qy      1061 TCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTG 1120
            || |||||| |||| ||||| |||||||||  |  ||| | ||||| || |||||  ||
Db       699 TCAAAGAACCGGGTAGAAGTCTGTTCTCTGTTACTTAGTCACGGTGCCGATCCAACCTTG 758

Qy      1121 CTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGA 1180
            |||||||  ||     |||||||||  | |||  |||||||  || ||   ||  |     ||||||
Db       759 GTCAATTGCCATGGCAAAAGTGCTGTCGACATGGCTCCAACTCCTGAGCTCCGGGAAAGA 818

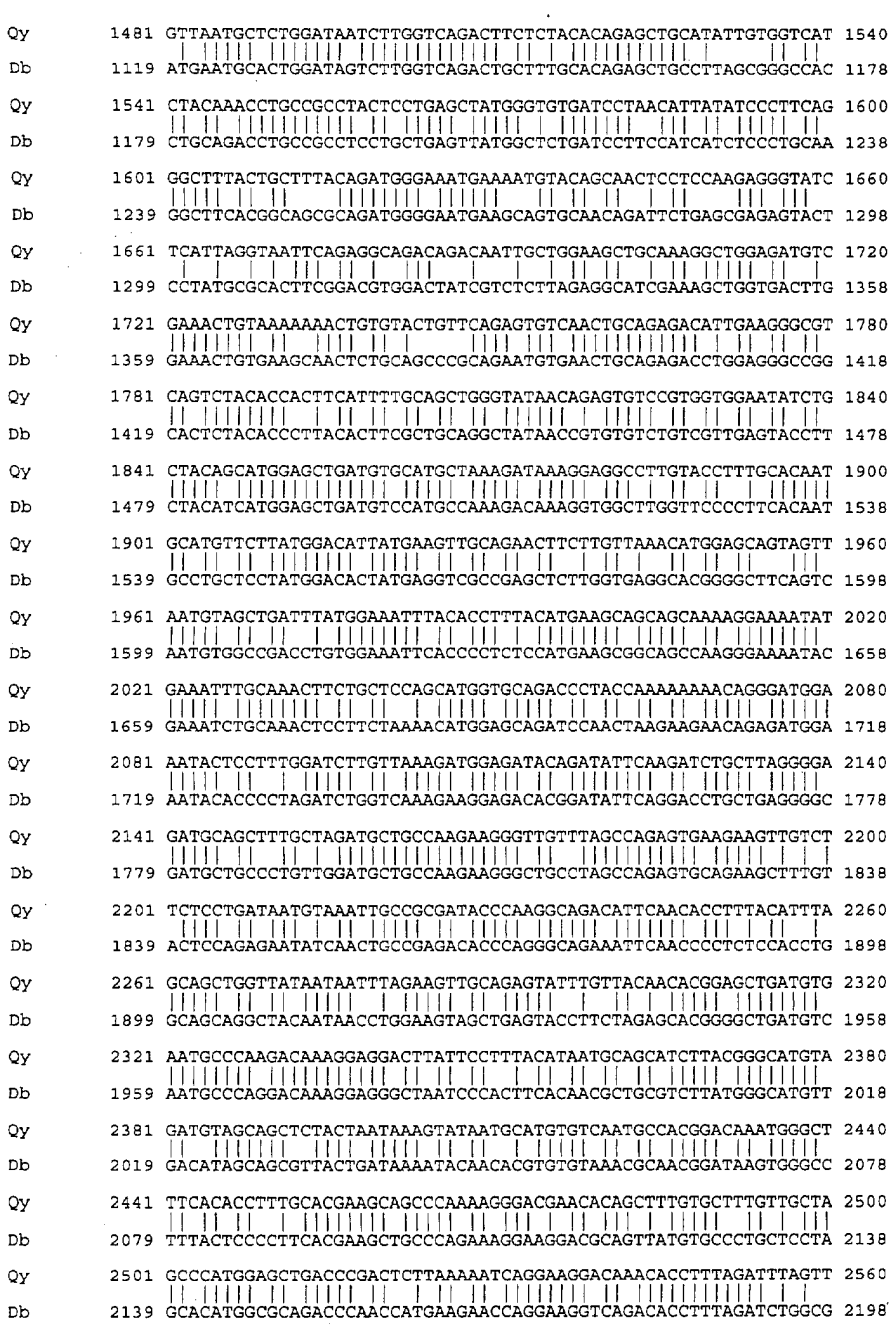
Qy      1181 TTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACT 1240
            |  |||||||||||||||||||||||||| || || || ||||| ||  |||||||| || |  |
Db       819 CTGACATATGAATTTAAAGGCCATTCTTTACTCCAAGCCGCCAGAGAAGCAGACCTGGCC 878

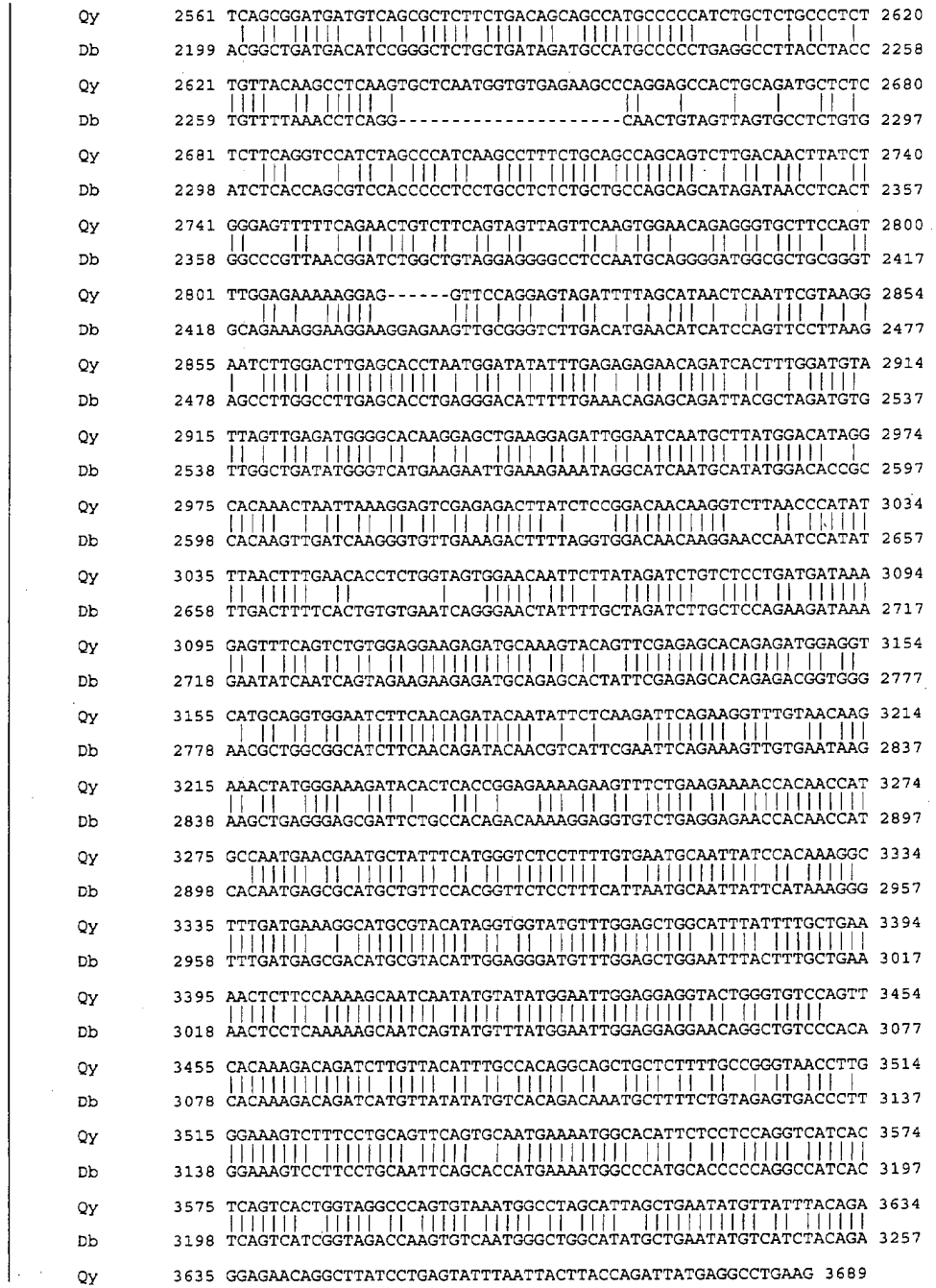
Qy      1241 CGAATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAA 1300
               | | ||||||   ||  || |||||||  | ||| |||||||| || ||  | ||||||
Db       879 AAAGTTAAAAAAAACGCTTGCTTTGGAAATCATTAACTTCAAGCAGCCACAGTCTCATGAA 938

Qy      1301 ACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTG 1360
            || |||||||| |||||||  || ||     |||||||| |||||||  |    || |||
Db       939 ACTGCATTGCACTGTGCTGTGGCCTCCTTGCATCCCAAACGAAAGCAGGTGGCAGAGCTG 998

Qy      1361 TTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCAC 1420
            || || || || |||||  ||| |||| |||||||  || || ||| |||| ||||||||
Db       999 TTACTGAGGAAGGGAGCCAACGTCAACGAAAAGAACAAGGATTTCATGACCCCTCTGCAT 1058

Qy      1421 GTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAG 1480
            || ||| |||| | |||||||||||||||  | |||||  ||   || || | |||||||
Db      1059 GTTGCAGCTGAAAGAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCACGGAGCAAAG 1118
```





```
                    || |||||||| ||||| || ||| | || || |||||||| |||| ||| ||||
Db       3258 GGGGAACAGGCATATCCAGAATATCTCATCACCTACCAGATCATGAAGCCAGAAG 3312
```

RESULT 2
AK047094
LOCUS AK047094 1967 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:B930018O21 product:TANKYRASE-RELATED PROTEIN (FRAGMENT) homolog [Homo sapiens], full insert sequence.
ACCESSION AK047094
VERSION AK047094.1 GI:26338577
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1967)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES Location/Qualifiers
source 1. .1967
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:B930018O21"
/db_xref="MGI:2412634"
/db_xref="taxon:10090"
/clone="B930018O21"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
CDS 166. .>1965
/note="unnamed protein product; TANKYRASE-RELATED PROTEIN (FRAGMENT) homolog [Homo sapiens] (SPTR|Q9HAS4, evidence: FASTY, 94.6%ID, 51.4%length, match=1959) putative"
/codon_start=1
/protein_id="BAC32960.1"
/db_xref="GI:26338578"
/translation="MSGRRCAGGGAACASAGAEAVEPSARELFEACRNGDVERVKRLV
TPEKVNSRDTAGRKSTPLHFAAGFGRKDVVEYLLQNGANVQARDDGGLIPLHNACSFG
HAEVVNLLLQHGAHPNARDNWNYTPLHEAAIKGKIDVCIVLLQHGAEPTIRNTDGRTA
LDLADPSAKAVLTGDYKEDELLESARSGNEEKMMALLTPLNVNCHASDGRKSTPLHLA
AEYNRVKIVQLLLHHGADVHAKDKGDLVPLHNACSYGHYEVTELLVKHGACVNAMDLW
QFTPLHEAASKNRIEVCSLLLSYGADPTLLNCHNKSAIDLAPTAQLKERLSYEFKGHS
LLQAAREADVTRIKKHLSLEMVNFKHPQTHETALHCAAASPYPKRKQICELLLRKGAN
TNEKTKEFLTPLHVASENAHNDVVEVVVKHEAKVNALDSLGQTSLHRAAHCGHLQTCR
LLLSYGCDPNIISLQGFTALQMGNENVQQLLQEGASLGHSEADRQLLEAAKAGDVETV
KKLCTVQSVNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDKGGLVPLHNA
CSYGHYEVAELLVKHGAVVNVADLWKFTPLHEAA"
ORIGIN

Query Match 36.1%; Score 1369.8; DB 11; Length 1967;
Best Local Similarity 90.8%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||| ||||| || |||||||||||| | |||||||| ||||||| ||| |||||
Db        363 AGGTTTTGGACGGAAGGATGTAGTTGAATATCTTCTTCAGAACGGTGCAAATGTGCAAGC 422

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              |||||||||||||| |||||||||||||||| |||||||||||||||||||||||||||
Db        423 GCGTGATGATGGGGGTCTTATTCCTCTTCATAACGCATGCTCTTTTGGTCATGCTGAAGT 482

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||| ||||| | ||| |||||| |||| |||||||| ||||||||||||||||
Db        483 AGTCAATCTTCTTTTACAACACGGTGCACACCCAAATGCTCGGGATAATTGGAATTATAC 542

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              |||||||||||||||||||||||||||||| |||||||||||||||||| ||||||||
Db        543 TCCTCTCCATGAAGCTGCAATTAAAGGAAAAATTGATGTTTGCATTGTGCTTTTACAGCA 602

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        603 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGACCC 662

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              |||||||||||| || || |||||||| |||||| |||||||||| | ||||||||||||
Db        663 ATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGGAAGATGAACTTCTGGAAAGTGCCAG 722

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
```

```
Db        723 GAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCCATGC 782
Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db        783 AAGTGATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGAATACAACAGAGTAAAAAT 842
Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db        843 TGTACAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATCTGGT 902
Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db        903 ACCACTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCA 962
Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db        963 CGGTGCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAGCTTC 1022
Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db       1023 TAAGAACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCAACCCTGCT 1082
Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db       1083 GAACTGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAAGATT 1142
Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db       1143 ATCGTATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTACAAG 1202
Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db       1203 AATCAAAAAACATCTTTCTCTGGAAATGGTGAATTTCAAACATCCTCAGACACATGAAAC 1262
Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db       1263 AGCACTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAGATATGTGAACTGCT 1322
Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db       1323 GCTCAGAAAGGGAGCAAACACCAACGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1382
Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db       1383 GGCTTCTGAGAACGCTCACAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1442
Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db       1443 TAATGCTTTGGATAGTCTTGGACAGACGTCATTACACAGAGCTGCACACTGTGGTCACCT 1502
Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db       1503 GCAGACCTGCCGCCTGCTCCTGAGCTATGGGTGTGATCCAAACATCATATCCCTTCAGGG 1562
Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db       1563 TTTCACCGCCTTGCAAATGGGAAATGAAAATGTGCAGCAGCTGCTCCAAGAGGGCGCCTC 1622
Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db       1623 ACTAGGTCACTCAGAGGCGGACAGACAGCTGCTGGAAGCTGCAAAGGCTGGTGACGTTGA 1682
Qy       1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db       1683 AACTGTAAAAAAACTCTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCA 1742
Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db       1743 GTCAACCCCACTCCACTTTGCAGCTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCT 1802
Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db       1803 GCAACATGGAGCTGATGTTCATGCTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGC 1862
Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db       1863 ATGTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAA 1922
Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGC 2007
Db       1923 TGTCGCTGACTTGTGGAAGTTCACACCTTTACATGAAGCTGCAGC 1967
```

RESULT 3
AK080913

```
LOCUS       AK080913               3194 bp    mRNA    linear   HTC 20-SEP-2003
DEFINITION  Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length
            enriched library, clone:B430206N15 product:TRF1-INTERACTING
            ANKYRIN-RELATED ADP-RIBOSE POLYMERASE homolog [Homo sapiens], full
            insert sequence.
ACCESSION   AK080913
VERSION     AK080913.1  GI:26099551
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3194)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
```

```
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .3194
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:B430206N15"
                     /db_xref="MGI:2412078"
                     /db_xref="taxon:10090"
                     /clone="B430206N15"
                     /sex="male"
                     /tissue_type="adipose"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="4 days neonate"
     misc_feature    27. .>3193
                     /note="TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE
                     POLYMERASE homolog [Homo sapiens] (SPTR|O95271, evidence:
                     FASTY, 95.9%ID, 80%length, match=3167)
                     putative"
ORIGIN

  Query Match            35.8%;  Score 1360.4;  DB 11;  Length 3194;
  Best Local Similarity  72.0%;  Pred. No. 0;
  Matches 1827;  Conservative    0;  Mismatches  681;  Indels   28;  Gaps    3;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              |||||||||  ||||  || ||  ||  ||  | |||||  ||||  |||||  | ||||||  ||
Db        678 AGGTTTTGGAAGGAAGGATGTTGTACAACACTTGCTGCAGATGGGTGCTAATGTCCACGC 737

Qy        463 A-CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAG 521
              | ||||||| |||||  ||  ||  || ||||||||||  ||||||||  ||  |||||||| |
Db        738 CCCGTGATGACGGGGGTCTCATCCCACTTCATAATGCCTGCTCTTTCGGCCATGCTGAGG 797

Qy        522 TAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATA 581
              | ||  |  |||   |  ||  |  ||  |||||  ||  |||||   |||||||  |||||  ||||
Db        798 TTGTGAGTCTGTTACTGTGTCAAGGAGCAGATCCGAATGCCAGAGATAACTGGAACTATA 857

Qy        582 CTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGC 641
              | ||  ||  ||||||||||  ||||||||  ||  ||||||||  ||  |||||||||  |  ||||
Db        858 CACCACTGCATGAAGCTGCCATTAAAGGGAAAATTGATGTCTGTATTGTGCTGCTGCAGC 917

Qy        642 ATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATC 701
              | |||||  ||  ||||  ||||||  ||  ||  |||||  |   ||||   ||||   |  |||||||
Db        918 ACGGAGCCGATCCAAACATCCGGAACACCGATGGGAAATCAGCTCTGGACCTGGCAGATC 977

Qy        702 CATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCA 761
              | ||  ||  |||||  ||  ||  ||  ||||||||  ||  |||||  ||||||  |||||   ||| |
Db        978 CTTCAGCAAAAGCTGTCCTCACAGGTGAATACAAAAAAGACGAACTCCTAGAAGCTGCTA 1037

Qy        762 GGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACG 821
              |||||||  ||||||||||||  |  ||||||  ||||  ||  ||  |||||||  ||  |||||  |
Db       1038 GGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCTCTAAATGTGAATTGCCATG 1097

Qy        822 CAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGA 881
              ||||||||||  |||||||  |||||  |||||  ||||||||||  ||  ||||||||   ||

Db       1098 CAAGTGATGGACGAAAGTCTACTCCGTTACACCTGGCAGCAGGCTACAACAGAGTTCGGA 1157

Qy        882 TTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGG 941
              | ||  ||||||  |  ||  ||  |||||  ||||||||||  ||  |||||  |||||||    || |
Db       1158 TAGTTCAGCTGCTTCTACAGCATGGTGCTGATGTCCACGCGAAAGACAAAGGTGGACTTG 1217

Qy        942 TACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGC 1001
              ||||  |  ||  |||||  ||  ||  ||  ||  |||||  ||  ||  |||||   ||  |  ||||
Db       1218 TACCTCTTCATAATGCATGCTCATACGGACACTATGAGGTCACAGAACTGCTGCTAAAGC 1277

Qy       1002 ATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTT 1061
              ||||  ||  |||||  ||  ||  |||||  |||||||  |||||  ||  ||  |||||  ||  ||||
Db       1278 ATGGAGCTTGTGTCAACGCTATGGATCTGTGGCAGTTCACACCGCTGCATGAAGCTGCTT 1337

Qy       1062 CTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGC 1121
              | ||||||  ||||  |||||  ||||||||  |  |  |||  |  |||||  ||  |||||   ||
Db       1338 CAAAGAACCGGGTAGAAGTCTGTTCTCTGTTACTTAGTCACGGTGCCGATCCAACCTTGG 1397

Qy       1122 TCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGAT 1181
              |||||||  ||  |||||||||  |  |||  |||||||  ||  ||  ||  |  ||||||
Db       1398 TCAATTGCCATGGCAAAAGTGCTGTCGACATGGCTCCAACTCCTGAGCTCCGGGAAAGAC 1457

Qy       1182 TAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTC 1241
              |  ||||||||||||||||||  ||  ||  |||||  ||  ||||||  ||  |  |
Db       1458 TGACATATGAATTTAAAGGCCATTCTTTACTCCAAGCCGCCAGAGAAGCAGACCTGGCCA 1517

Qy       1242 GAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAA 1301
              | | ||||||   ||  ||  |||||||  | ||  ||||||||  ||  ||  | |||||||
Db       1518 AAGTTAAAAAAACGCTTGCTTTGGAAATCATTAACTTCAAGCAGCCACAGTCTCATGAAA 1577

Qy       1302 CAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGT 1361
              | ||||||| |||||||  ||  ||  |||||||  |||||||  |  ||  ||||||||
Db       1578 CTGCATTGCACTGTGCTGTGGCCTCCTTGCATCCCAAACGAAAGCAGGTGGCAGAGCTGT 1637

Qy       1362 TGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACG 1421
              | ||  ||  ||  |||||  |||  ||||  |||||||  ||  ||  |||  ||||  |||||||  |
Db       1638 TACTGAGGAAGGGAGCCAACGTCAACGAAAAGAACAAGGATTTCATGACCCCTCTGCATG 1697

Qy       1422 TGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGG 1481
              | |||  ||||  |  ||||||||||||||||  |  |||||  ||  ||  ||  |  |||||||
Db       1698 TTGCAGCTGAAAGAGCTCATAATGATGTCATGGAAGTTCTGCATAAGCACGGAGCAAAGA 1757

Qy       1482 TTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATC 1541
              | |||||  |||||||  ||||||||||||  ||  |  ||||||||||  |  ||  ||  |
Db       1758 TGAATGCACTGGATAGTCTTGGTCAGACTGCTTTGCACAGAGCTGCCTTAGCGGGCCACC 1817

Qy       1542 TACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGG 1601
              | ||  ||||||||||||  ||  |||||  |||||  |  |||||||  |||  ||  |||||  ||  |
Db       1818 TGCAGACCTGCCGCCTCCTGCTGAGTTATGGCTCTGATCCTTCCATCATCTCCCTGCAAG 1877

Qy       1602 GCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCT 1661
              ||||  ||  ||  ||||||||  |||||||  ||  ||  ||  |  ||  |||  |||
Db       1878 GCTTCACGGCAGCGCAGATGGGGAATGAAGCAGTGCAACAGATTCTGAGCGAGAGTACTC 1937

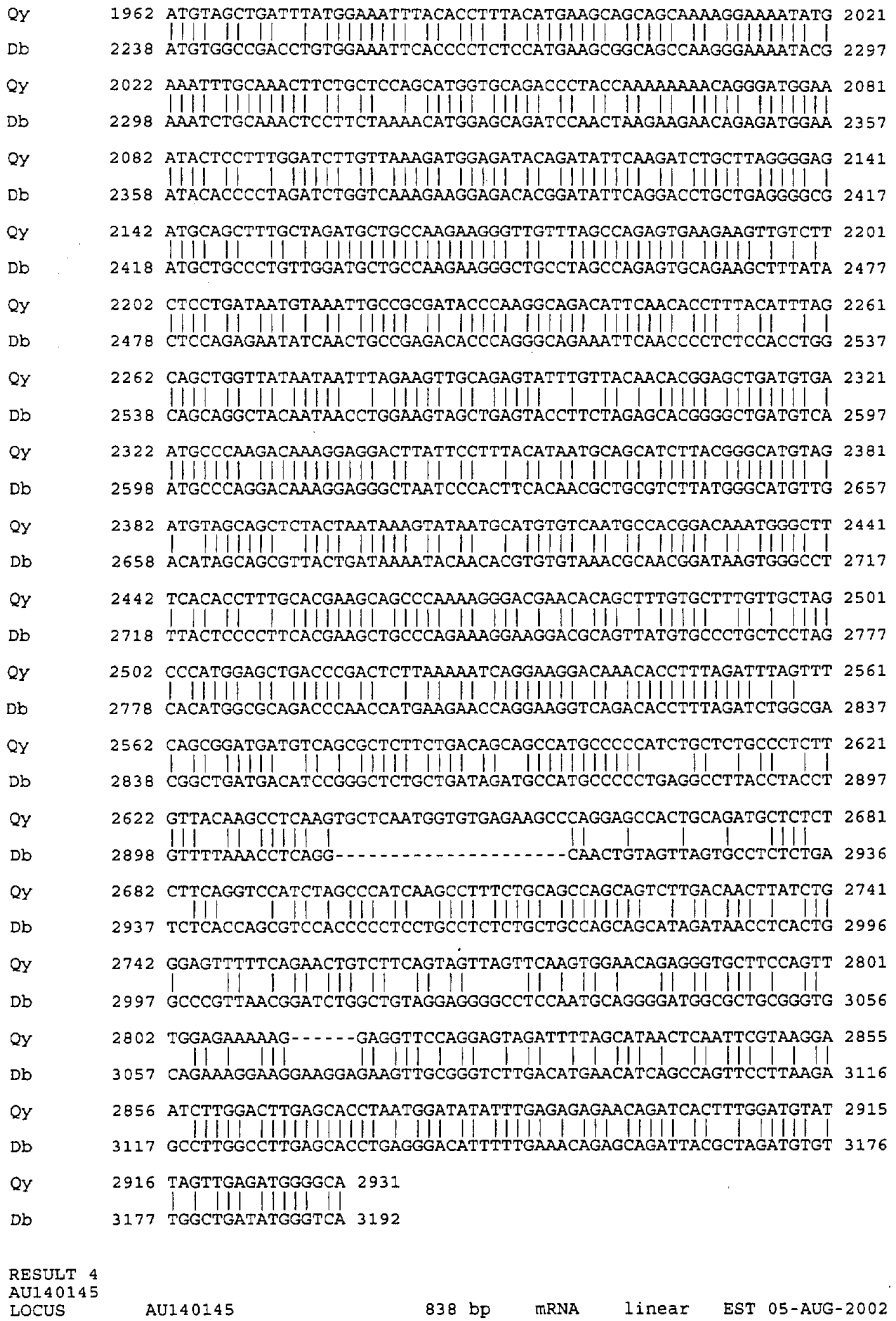
Qy       1662 CATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCG 1721
              | |  | |  |||  ||  |  |||  |  |  ||  ||  |||  |||||  ||  |  |
Db       1938 CTATGCGCACTTCGGACGTGGACTATCGTCTCTTAGAGGCATCGAAAGCTGGTGACTTGG 1997

Qy       1722 AAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTC 1781
              |||||||  ||  ||||  ||  |  ||||  |||  ||||||||||||  |  ||  ||  |
Db       1998 AAACTGTGAAGCAACTCTGCAGCCCGCAGAATGTGAACTGCAGAGACCTGGAGGGCCGGC 2057

Qy       1782 AGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGC 1841
              | ||||||||  |  ||  ||  ||  ||  ||||||  |  |||||  ||  ||  ||  ||  |
Db       2058 ACTCTACACCCTTACACTTCGCTGCAGGCTATAACCGTGTGTCTGTCGTTGAGTACCTTC 2117

Qy       1842 TACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATG 1901
              ||||  ||||||||||||||  |||||  |||||  |||||  |||  |  ||  ||  |  |||||||
Db       2118 TACATCATGGAGCTGATGTCCATGCCAAAGACAAAGGTGGCTTGGTTCCCCTTCACAATG 2177

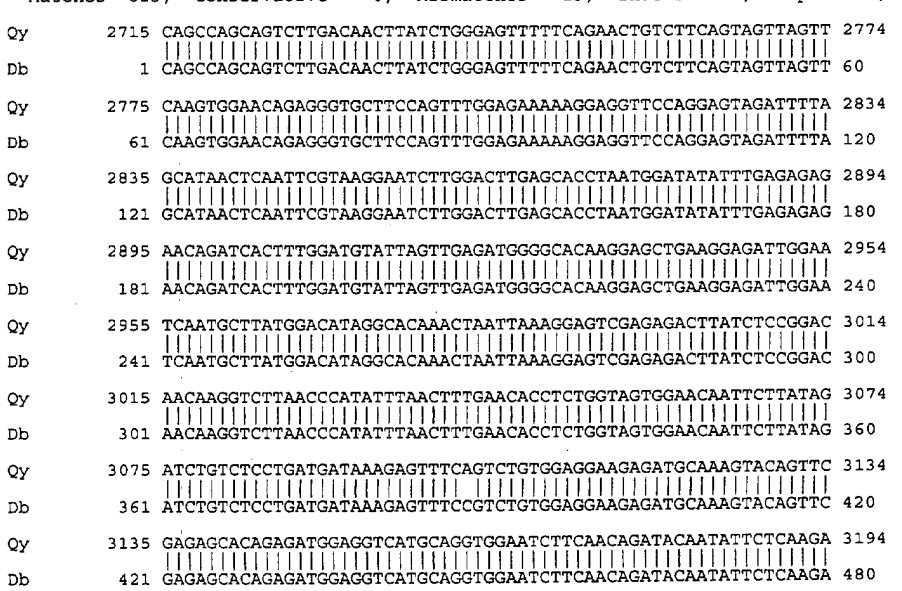
Qy       1902 CATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTA 1961
              | ||  ||  ||||||||  |||||  ||  ||  ||  |||  |  ||  ||  ||||  |
Db       2178 CCTGCTCCTATGGACACTATGAGGTCGCCGAGCTCTTGGTGAGGCACGGGGCTTCAGTCA 2237
```



RESULT 4
AU140145
LOCUS AU140145 838 bp mRNA linear EST 05-AUG-2002

DEFINITION AU140145 PLACE2 Homo sapiens cDNA clone PLACE2000021 5', mRNA sequence.
ACCESSION AU140145
VERSION AU140145.1 GI:11001666
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 838)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES Location/Qualifiers
source 1..838
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE2000021"
/tissue_type="placenta"
/clone_lib="PLACE2"
/note="Vector: pME18SFL3"
ORIGIN

Query Match 20.7%; Score 784.2; DB 9; Length 838;
Best Local Similarity 98.0%; Pred. No. 7e-185;
Matches 823; Conservative 0; Mismatches 13; Indels 4; Gaps 3;



```
Qy       3195 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 540

Qy       3255 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3314
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 600

Qy       3315 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3374
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 660

Qy       3375 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3434
             |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        661 CTGGCATTTATTTTGCTGAAAACTCTT-CAAAAGCAATCAATATGTATATGGAATTGGAN 719

Qy       3435 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3494
             || || |||||||||||||||||||||| ||||||| ||||||||||||| |||||||||||
Db        720 GANGT-CTGGGTGTCCAGTTCACAAAGACCGATCTTGGTACATTTGCCACANGCAGCTGC 778

Qy       3495 TCTTTTG--CCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 3552
             |||||||  | ||  | ||||||||||||||| |||||||||| ||||||| |||||||
Db        779 TCTTTTGGCCGGGTAACCCTTGGGAAAGTCTTTTCTGCAGTTCANTGCAATGNAAATGGC 838
```

RESULT 5
CB183123
LOCUS CB183123 955 bp mRNA linear EST 31-JAN-2003
DEFINITION AGENCOURT_11385409 NIH_MGC_164 Mus musculus cDNA clone IMAGE:30244272 5', mRNA sequence.
ACCESSION CB183123
VERSION CB183123.1 GI:28181571
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 955)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0316 row: o column: 01
High quality sequence stop: 779.
FEATURES Location/Qualifiers
source 1. .955
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30244272"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_164"
/note="Vector: pCMV-SPORT6.1; Site_1: EcoRV; Site_2: NotI; Non-normalized full-length enriched library from pooled mouse embryonic limb, maxilla and mandible, day 10.5 and 11.5 (size selected for the 0.5-1 kb fragments) Cloned directionally, priming method: Oligo-dT. cDNA enrichment: >1k bp, Average insert size 1.8k bp. Priming sequence: 5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T) 3'. Tissue contributed by, David Rowe. Library constructed by ResGen, Invitrogen Corp."
ORIGIN

Query Match 18.1%; Score 687.8; DB 14; Length 955;
Best Local Similarity 88.5%; Pred. No. 9.5e-161;
Matches 768; Conservative 0; Mismatches 98; Indels 2; Gaps 2;

```
Qy       2816 GTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTA 2875
             | | |||| | |||||||| |||||||| ||| |||||||| ||||||||||||||| |
Db          1 GATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACTTG 60

Qy       2876 ATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAG 2935
             |||||||||||||||||||||||||| || || |||| |||||| |||| ||||||
Db         61 ATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAAG 120

Qy       2936 GAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTC 2995
             || |||| |||||||||||||||||||| |||||||| | |||| || ||||||||||
Db        121 GAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACAAGCTGATTAAAGGAGTT 180

Qy       2996 GAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGT 3055
             || || || |||| |||||||||||||||| |||||||||| | |||| |||||||
Db        181 GAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAACTCTAAACAACTCTGGT 240

Qy       3056 AGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAA 3115
             |||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||| |||
Db        241 AGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAA 300

Qy       3116 GAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAAC 3175
             ||||||| ||||| || || |||||||||||||| |||| ||||| || |||||||
Db        301 GAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAAC 360

Qy       3176 AGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACT 3235
             |||||||| |||||||||||||||||||||||||||||| | |||||||||| ||
Db        361 AGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTGGGAAAGATATACA 420

Qy       3236 CACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTT 3295
             |||||||||||||||||||||||||||||||||||||||||||||| ||| | ||| |||||
Db        421 CACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCAAACGAAAGGATGTTATTT 480

Qy       3296 CATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTAC 3355
             |||||||||||||||||||||| |||||||| || |||||||||||||||||||| |||
Db        481 CATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTGATGAAAGGCATGCATAC 540

Qy       3356 ATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAA 3415
             |||||||| |||||||||||| |||||||||||| |||||||||||||||||||||||||
Db        541 ATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTCCAAAAGCAATCAA 600

Qy       3416 TATGTATAT-GGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTA 3474
             ||||| ||| |||||||||| || || || |||||| |||||||||||||||| |||||
Db        601 TATGTGTATNGGAATTGGAGGTGGCACCGGATGTCCAATTCACAAAGACAGATCGTGTTA 660

Qy       3475 CATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTT 3534
             |||||| |||||||||||||| ||||| || |||||||||||||||||||| |||||||
Db        661 CATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAAAGTCTTTCTTGCAGTT 720

Qy       3535 CAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAG 3594
             ||| |||||||||||||||||||||||||| || |||||||| |||||||| |||||||
Db        721 CAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGGTCACTGGCCGGCCCAG 780

Qy       3595 TGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGA 3654
             ||||||| |||||||||||||||||||||||||||||||||||||||||||   |||||
Db        781 TGTAAAT-GCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTATCCCTGA 839

Qy       3655 GTATTTAATTACTTACCAGATTATGAGG 3682
                 | | ||| |  || | | || |
Db        840 TATCTATACTACCAATGAGCTGANGATG 867
```

RESULT 6
CB128566
LOCUS CB128566 675 bp mRNA linear EST 29-JAN-2003
DEFINITION K-EST0178055 L11SNU354 Homo sapiens cDNA clone L11SNU354-7-A10 5', mRNA sequence.
ACCESSION CB128566
VERSION CB128566.1 GI:28091690
KEYWORDS EST.
SOURCE Homo sapiens (human)

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 7  row: A  column: 10
            High quality sequence stop: 675.
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L11SNU354-7-A10"
                     /sex="M"
                     /tissue_type="Liver"
                     /cell_type="Polygonal"
                     /cell_line="SNU-354"
                     /lab_host="Top10F'"
                     /clone_lib="L11SNU354"
                     /note="Organ: Liver; Vector: pCNS-D2; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including
                     EcoRI site by treatment of T4 RNA ligase and the first
                     strand cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

  Query Match           17.7%;  Score 672.4;  DB 14;  Length 675;
  Best Local Similarity   99.7%;  Pred. No. 6.4e-157;
  Matches  673;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2417 GTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGA 2476
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGA 60

Qy       2477 ACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAA 2536
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAA 120

Qy       2537 GGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCC 2596
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        121 GGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCC 180

Qy       2597 ATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGC 2656
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGC 240

Qy       2657 CCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCA 2716
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCA 300

Qy       2717 GCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCA 2776
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCA 360

Qy       2777 AGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGC 2836
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGC 420

Qy       2837 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 2896
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 480

Qy       2897 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 2956
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 540

Qy       2957 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 3016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 600

Qy       3017 CAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 3076
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        601 CAAGGTCTTAACCCATATTTAACTNTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 660

Qy       3077 CTGTCTCCTGATGAT 3091
             ||||||||||||||||
Db        661 CTGTCTCCTGATGAT 675


RESULT 7
BM457025
LOCUS       BM457025                1128 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6406851 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583257
            5', mRNA sequence.
ACCESSION   BM457025
VERSION     BM457025.1  GI:18506065
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1128)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12346  row: a  column: 18
            High quality sequence stop: 507.
FEATURES             Location/Qualifiers
     source          1. .1128
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5583257"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_92"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           17.7%;  Score 672.2;  DB 12;  Length 1128;
  Best Local Similarity   97.7%;  Pred. No. 7.9e-157;
```

Matches 724; Conservative 0; Mismatches 13; Indels 4; Gaps 4;

```
Qy       1957 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 2016
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAA 60

Qy       2017 ATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGA 2076
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGA 120

Qy       2077 TGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 2136
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAG 180

Qy       2137 GGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTT 2196
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTT 240

Qy       2197 GTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACA 2256
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACA 300

Qy       2257 TTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGA 2316
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGA 360

Qy       2317 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCA 2376
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCA 420

Qy       2377 TGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATG 2436
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATG 480

Qy       2437 GGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTT 2496
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTT 540

Qy       2497 GCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA-CCTTTAGATT 2555
             ||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db        541 GCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCCTTTACATT 600

Qy       2556 TAGTTTCAGCGGATGATGTCAGCG-CTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTG 2614
             |||||||||||||||||||||||| |||||||| |||||||| |||||||||||||||||||
Db        601 TAGTTTCAGCGGATGATGTCAGCGCCTCTTCTGACACCAGCCATGGCCCCATCTGCTCTG 660

Qy       2615 CCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTG-TGAGAAGCCCAGGAGCCACTGCAGA 2673
             |||||||| ||| ||||||||| |||| |||||| |||||||||  ||||||||||||||||
Db        661 CCCTCTTGGTACCAGCCTCAAGCGCTCCATGGTGCTGAGAAGCCACGGAGCCACTGCAGA 720

Qy       2674 TG-CTCTCTCTTCAGGTCCAT 2693
             || || |||||||||||| | |
Db        721 TGCCTTTCTCTTCAGGTTCCT 741
```

RESULT 8
CB170279/c

```
LOCUS       CB170279                754 bp    mRNA    linear   EST 30-JAN-2003
DEFINITION  CEY603161575.R1 CSEQFXN42 testes Bos taurus cDNA, mRNA sequence.
ACCESSION   CB170279
VERSION     CB170279.1  GI:28156407
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 754)
  AUTHORS   Adelson,D.L. and Gill,C.A.
  TITLE     Bovine ESTs (Adelson and Gill)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: David L. Adelson
            Animal Breeding and Genetics
            Texas A&M University
            Animal Science Dept., TAMU-2471, College Station, TX 77843-2471,
            USA
            Tel: 9798452616
            Fax: 9798456970
            Email: david.adelson@tamu.edu.
FEATURES             Location/Qualifiers
     source          1. .754
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="testes"
                     /clone_lib="CSEQFXN42 testes"
                     /note="Organ: testes; Vector: pBluescript SK+; Site_1:
                     NotI; Site_2: EcoRI; sequence 5' of the insert
                     (5'-NNN. . .NNNinsert)
                     GCGAATTGGAGCTCCACCGCGGTGGCGGCCGCGGCTCGAG. Sequence 3' of
                     the inserts (AAGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAG.
                     normalized Rd 2 library, sequenced 3' with M13R primer."
ORIGIN
```

Query Match 17.5%; Score 665.2; DB 14; Length 754;
Best Local Similarity 93.5%; Pred. No. 4.1e-155;
Matches 705; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
Qy        591 ATGAAGCTGCAATTAAAGGAAAGATTGATGTTTG-CATTGTGCTGTTACAGCATGGAGCT 649
              |||||||||||||||||||||||||||||||||| || ||||||||| ||||||||||||||
Db        754 ATGAAGCTGCAATTAAAGGAAAGATTGATGTCTGCCATTGTGCTCTTACAGCATGGAGCT 695

Qy        650 GAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCC 709
              |||||||||| |||||||||||||| |||||||||||||||||||| || ||||||
Db        694 GAGCCAACCATTCGAAATACAGATGGACGGACAGCATTGGATTTAGCAGACCCGTCTGCC 635

Qy        710 AAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGC 769
              |||||||||| |||||||||||||||| |||||||||| |||||||||||||||||||||
Db        634 AAAGCAGTGCTGACTGGTGAATATAAGAAGGATGAACTCTTGGAAAGTGCCAGGAGTGGC 575

Qy        770 AATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGAT 829
              ||||| |||||||||||||||| || |||||||||||||| ||||| || |||||||||
Db        574 AATGAGGAAAAAATGATGGCTCTGCTTACACCATTAAATGTAAACTGTCATGCAAGTGAT 515

Qy        830 GGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAG 889
              |||||||||||||||||| |||||||||||| || ||||||||||||||||||||||||
Db        514 GGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGTTACAACAGAGTAAAGATTGTACAG 455

Qy        890 CTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTA 949
              |||||  |||||||||||||||||||||||||||||||||||||||||||||||   |
Db        454 CTGTTGTTGCAACATGGAGCTGATGTTCATGCTAAAGATAAAGGTGATCTGGTACCCCTG 395

Qy        950 CACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCC 1009
              ||||| |||||||| |||||||||||||||||||||||||||| |||||||||||||||
Db        394 CACAACGCCTGTTCCTATGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCATGGTGCC 335

Qy       1010 TGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAAC 1069
              || |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db        334 TGCGTAAATGCAATGGACTTGTGGCAGTTCACTCCTCTCCATGAGGCTGCTTCCAAGAAC 275

Qy       1070 AGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGT 1129
              |||||||||| |||||||||||| | |||||||| ||||| |||||||||||  ||| |||
Db        274 AGGGTTGAAGTGTGTTCTCTTCTCCTGAGTTATGGCGCAGATCCAACACTGCTGAACTGT 215

Qy       1130 CACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATAT 1189
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        214 CACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGGTTAGCATAT 155

Qy       1190 GAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAA 1249
              |||||||||||||||| ||||| || |||||||| |||||||||||||| |||||||||
Db        154 GAATTTAAAGGCCACTCCTTGCTTCAGGCTGCACGGGAAGCTGATGTTACACGAATCAAA 95

Qy       1250 AAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTG 1309
              |||||||||||||||||||||||| || |||||||||||||||||||||||||||||||
Db         94 AAACATCTCTCTCTGGAAATGGTGAACTTTAAGCATCCTCAAACACATGAAACAGCATTG 35
```

```
Qy       1310 CATTGTGCTGCTGCATCTCCATATCCCAAAAGAA 1343
              |||||||||||||||||||||||||||||||| ||
Db         34 CATTGTGCTGCTGCATCTCCATATCCCAAAAAAA 1
```

```
RESULT 9
CF407574/c
LOCUS       CF407574                923 bp    mRNA    linear   EST 02-SEP-2003
DEFINITION  CH3#049_C04T3 Canine heart normalized cDNA Library in pBluescript
            Canis familiaris cDNA clone CH3#049_C04 3', mRNA sequence.
ACCESSION   CF407574
VERSION     CF407574.1  GI:34408600
KEYWORDS    EST.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 923)
  AUTHORS   Yi,Y., Desai,R., Olarte,M., Henthorn,P. and George A.L.
  TITLE     Expressed sequence tags from Canine heart
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: CH3#049_C04T7
            Contact: George AL
            Division of Genetic Medicine
            Vanderbilt University
            529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
            Tel: 615 936 2660
            Fax: 615 936 2661
            Email: al.george@vanderbilt.edu
            Insert Length: 2028   Std Error: 0.00
            Seq primer: T3: ATTAACCCTCACTAAAGGGA
            High quality sequence start: 65
            High quality sequence stop: 838.
FEATURES             Location/Qualifiers
     source          1. .923
                     /organism="Canis familiaris"
                     /mol_type="mRNA"
                     /db_xref="taxon:9615"
                     /clone="CH3#049_C04"
                     /tissue_type="heart"
                     /cell_type="heart"
                     /dev_stage="mixed developmental stages (adult, 30 day - 40
                     day fetal)"
                     /clone_lib="Canine heart normalized cDNA Library in
                     pBluescript"
                     /note="Organ: heart; Vector: pBluescript; Site_1: 5' of
                     vector NotI; Site_2: 3' of vector EcoRI; Tissue source:
                     dog heart (adult, 30 day - 40 day fetal), right and left
                     atria and ventricle. Dog breed - mixed (beagle, German
                     shepherd,  pointer, Irish setter). Library construction:
                     oligo-dT primed"
ORIGIN
```

```
  Query Match             17.2%;  Score 652;  DB 14;  Length 923;
  Best Local Similarity   94.3%;  Pred. No. 8.6e-152;
  Matches  676;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy       2584 TCTGACAGCAGCCATGCCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAA 2643
              |||||||||||||||||||| | ||||||||| || ||||| || ||||||||||||||
Db        893 TCTGACAGCAGCCATGCCCCCTTNTGCTCTGCCTTCATGTTATAAACCTCAAGTGCTCAA 834

Qy       2644 TGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATC 2703
              ||| |||||||||||||||| || |  |||||||| |||||||||||||| ||||||||
Db        833 TGGCGTGAGAAGCCCAGGAGCTACCGGCGATGCTCTGTCTTCAGGTCCATCCAGCCCATC 774

Qy       2704 AAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTC 2763
              ||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        773 AAGTCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCTGAACTGTCTTC 714

Qy       2764 AGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGG 2823
              |||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db        713 AGTAGTTAGTTCAAGTGGAACAGAGGGTGCCTCTAGTTTGGAGAAAAAGGAGGTTCCAGC 654
```

```
Qy       2824 AGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATAT 2883
              |||||||||||||||||||| || |||||||||||||| |||||||||| ||||||||
Db        653 AGTAGATTTTAGCATAACTCAGTTTGTAAGGAATCTTGGGCTTGAGCACCTGATGGATAT 594

Qy       2884 ATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAA 2943
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        593 ATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTGGAGATGGGGCACAAGGAGCTGAA 534

Qy       2944 GGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACT 3003
              |||||||||||||| |||||||||||||| |||||||||||||||||||| ||||||||
Db        533 GGAGATTGGAATCAACGCTTATGGACATAGACACAAACTAATTAAAGGAGTTGAGAGACT 474

Qy       3004 TATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAAC 3063
              ||| || |||||||||||||||||||||||| |||||| |||||||||||||| ||||||||
Db        473 TATATCTGGACAACAAGGTCTTAACCCATACTTAACTCTGAACACCTCTGGAAGTGGAAC 414

Qy       3064 AATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCA 3123
              |||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        413 AATTCTCATAGATCTATCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCA 354

Qy       3124 AAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAA 3183
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 AAGTACAGTCCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAA 294

Qy       3184 TATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAG 3243
              |||||||||||||||| |||||||||||||| |||||||||||| || ||| ||||
Db        293 TATTCTCAAGATTCAGAAAGTTTGTAACAAGAAATTATGGGAAAGATATACACACAGGAG 234

Qy       3244 AAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        233 AAAAGAAGTTTCTGAAGAAAACCACAACCATGCAAATGAAAGAATGCTATTTCATGG 177
```

```
RESULT 10
R64714
LOCUS       R64714                  659 bp    mRNA    linear   EST 12-JUN-1996
DEFINITION  EST54a09 WATM1 Homo sapiens cDNA clone 54a09 similar to BLASTX
            homology with sp|Q01485|ANKC_HUMAN ANKYRIN, BRAIN VARIANT 2, mRNA
            sequence.
ACCESSION   R64714
VERSION     R64714.1  GI:837352
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Bouillaud,F.
  TITLE     Study of expressed sequences tags in adipose tissue 1995
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Frederic Bouillaud
            Centre de Recherche sur l'Endocrinologie moleculaire et le
            Developpement
            CNRS
            9, Rue Jules Hetzel, Meudon Bellevue, 92190 France
            Tel: 33 1 45 07 52 87
            Fax: 33 1 45 07 58 90
            Email: bouillau@infobiogen.fr
            automatic cycle sequencing of PCR amplified insert, open reading
            frame 3. .659.
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /strain="Caucasian"
                     /db_xref="taxon:9606"
                     /clone="54a09"
                     /clone_lib="WATM1"
                     /note="Vector: lambda gt11; Site_1: EcoRI; White adipose
                     tissue, sub cutaneous, adult, female. Purification of
                     polyA mRNA, first strand priming with random
                     oligonucleotides. Ligation of adaptors EcoRI/NotI, cloning
                     in EcoRI site of lambda gt11."
```

ORIGIN

Query Match 17.1%; Score 649.2; DB 14; Length 659;
Best Local Similarity 98.6%; Pred. No. 4.1e-151;
Matches 649; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1416 TGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAG 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAG 60

Qy       1476 CAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTG 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db         61 CAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATHTTGTG 120

Qy       1536 GTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCC 1595
             |:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GYCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCC 180

Qy       1596 TTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGG 1655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGG 240

Qy       1656 GTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAG 1715
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        241 GTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTSCAAAGGCTGGAG 300

Qy       1716 ATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAG 1775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAG 360

Qy       1776 GGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAT 1835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAAT 420

Qy       1836 ATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGC 1895
             ||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATCTGCYACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGC 480

Qy       1896 ACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAG 1955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAG 540

Qy       1956 TAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAA 2015
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        541 TAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCCAAAGGGA 600

Qy       2016 AATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAG 2073
             ||:|||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db        601 AAWATGAAATTTGCCAACTTCTGCTCCAGCATGGTGCAGACCCCACCAAAAAAAACAG 658
```

RESULT 11
BG391376
LOCUS BG391376 691 bp mRNA linear EST 12-MAR-2001
DEFINITION 602417447F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4537138 5', mRNA sequence.
ACCESSION BG391376
VERSION BG391376.1 GI:13284824
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 691)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10461 row: m column: 11
High quality sequence stop: 684.

FEATURES Location/Qualifiers
source 1. .691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4537138"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 17.0%; Score 643.8; DB 12; Length 691;
Best Local Similarity 99.3%; Pred. No. 9.2e-150;
Matches 678; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

```
Qy       1706 AAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1765
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          10 AAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 69

Qy       1766 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1825
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          70 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 129

Qy       1826 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1885
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 189

Qy       1886 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1945
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 249

Qy       1946 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2005
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 309

Qy       2006 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCA-TGGTGCAGACCCTACCAA 2064
             |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||| ||
Db         310 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCACTGGTGCAGACCCTACAAA 369

Qy       2065 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 2124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         370 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 429

Qy       2125 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 2184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         430 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 489

Qy       2185 AGTGAAGAAGTTGTCTTCTCCTGATA-ATGTAAATTGCCGCGATACCCAAGGCAGACATT 2243
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db         490 AGTGAAGAAGTTGTCTTCTCCTGATACATGTAAATTGCCGCGATACCCAAGGCAGACATT 549

Qy       2244 CAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTAC 2303
             |||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db         550 CAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTA-TAGTTAC 608

Qy       2304 AACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 2363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         609 AACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAG 668

Qy       2364 CATCTTACGGGCATGTAGATGTA 2386
             |||||||||||||||||||||||
Db         669 CATCTTACGGGCATGTAGATGTA 691
```

```
RESULT 12
AL601027
LOCUS       AL601027                759 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp313G2239_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
            DKFZp313G2239 5', mRNA sequence.
ACCESSION   AL601027
VERSION     AL601027.1  GI:15164533
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Duesterhoeft, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp313G2239) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313G2239"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="313 (synonym: hlcc2)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match            16.9%;  Score 642.8;  DB 9;  Length 759;
  Best Local Similarity  97.2%;  Pred. No. 1.7e-149;
  Matches  669;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    2;

Qy       2374 GCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAA 2433
             ||||||| |||||||||||||||||||||||  |||||||||||||||||||||||| ||| 
Db         72 GCATGTANATGTAGCAGCTCTACTAATAAAGNNTAATGCATGTGTCAATGCCACGGGCAA 131

Qy       2434 ATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTT 2493
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||| 
Db        132 ATGGGCTTTCACACCTTTGCACGAAGCAGCCCAGAAGGGACGAACACAGCTTTGTGCTTT 191

Qy       2494 GTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGA 2553
              ||   | |||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        192 NTTTTTTGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGA 251

Qy       2554 TTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCT 2613
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        252 TTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCT 311

Qy       2614 GCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGA 2673
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        312 GCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGA 371

Qy       2674 TGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAA 2733
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||| 
Db        372 TGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGNCTTGACAA 431

Qy       2734 CTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGC 2793
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        432 CTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGC 491

Qy       2794 TTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAG 2853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| 
Db        492 TTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGNAAG 551

Qy       2854 GAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 2913
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        552 GAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGT 611

Qy       2914 ATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAG 2973
              |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        612 ANNAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAG 671

Qy       2974 GCACAAACTAATTAAA-GGAGTCGAGAGACTTATCTCCGGACAACAA-GGTCTTAACCCA 3031
             |||||||||||||||| |||||||||||||| |||||||||||||| || |||||||||  
Db        672 GCACAAACTAATTAAAGGGAGTCGAGAGACTNATCTCCGGACAACAAGGGGCTTAACCCA 731

Qy       3032 TATTTAACTTTGAACACCTCTGGTAGTG 3059
             |||||||||||||||||||||| || | 
Db        732 TATTTAACTTTGAACACCTCTGGNAGAG 759


RESULT 13
BQ885764
LOCUS       BQ885764                950 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8746924 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6392888
            5', mRNA sequence.
ACCESSION   BQ885764
VERSION     BQ885764.1  GI:22277782
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 950)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
             cDNA Library Preparation: ResGen, Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13883  row: h  column: 09
            High quality sequence stop: 641.
FEATURES             Location/Qualifiers
     source          1. .950
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6392888"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_130"
                     /note="Organ: otocysts; Vector: pCMV-SPORT6.1; Site_1:
                     EcoRV; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  Average insert size 1.95  kb.  Constructed by
                     ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            16.7%;  Score 634.4;  DB 13;  Length 950;
  Best Local Similarity  86.4%;  Pred. No. 2.2e-147;
  Matches  759;  Conservative    0;  Mismatches  111;  Indels    8;  Gaps    5;

Qy       2481 AGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGAC 2540
             |||||||||||||| ||| | |||||||||||||| || |||||| ||||||| ||||||| 
Db          1 AGCTTTGTGCTTTATTGTTGGCCCATGGAGCTGATCCTACTCTT-AAAATCAAGAAGGAC 59
```

```
Qy       2541 AAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGC 2600
Db         60 AAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGC 119

Qy       2601 CCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAG 2660
Db        120 CCCCCTCTGCTCTGCCTACGTGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCG 179

Qy       2661 GAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCA 2720
Db        180 GAGCCACTGCAGATGCTCTGTCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCA 239

Qy       2721 GCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTG 2780
Db        240 GCAGCCTCGACAACTTATCTGGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTG 299

Qy       2781 GAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAA 2840
Db        300 CAGCAGAAGGTGCTACTGGTTTGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAA 359

Qy       2841 CTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGA 2900
Db        360 CTCAGTTCATAAGGAACCTTGGACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGA 419

Qy       2901 TCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATG 2960
Db        420 TTACCTTAGATGTCCTAGTTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATG 479

Qy       2961 CTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAG 3020
Db        480 CCTATGGACATCGACACAAGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAACAAG 539

Qy       3021 GTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGT 3080
Db        540 GTCTTAATCCATATTTAACTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGT 599

Qy       3081 CTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGC 3140
Db        600 CTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGC 659

Qy       3141 ACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGA 3200
Db        660 ACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGA 719

Qy       3201 AGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG-AGAAAAGAAGTTTCTGAA 3259
Db        720 AGGTTTGTAACAAGAAATTGTGGGAAAGATATACACACCGGNAGAAAAGAAGTTTCTGAA 779

Qy       3260 GAAAACCACAACCATGC-CAATGAACGAATGCTATTTCATGGGTCT---CCTTTTGTGAA 3315
Db        780 GAAAACCACAACCATGCAAAACGAAAGGATGTTATTTTCATGGGGTCTCCTTTTTGTGAA 839

Qy       3316 TGCAATTATCCACAAAGG--CTTTGATGAAAGGCATGC 3351
Db        840 TGCGATTATCCATAAAGGGCTTTTGATGAAAGGCATGC 877
```

RESULT 14
CF539235

```
LOCUS       CF539235                743 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-GH0-chd-h-19-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
            IMAGE:30535386 5', mRNA sequence.
ACCESSION   CF539235
VERSION     CF539235.1  GI:34591604
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30535386"
                     /tissue_type="Whole brain"
                     /dev_stage="1, 5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GH0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     BonaTdo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN
```

Query Match 16.7%; Score 632.6; DB 14; Length 743;
Best Local Similarity 90.7%; Pred. No. 5.9e-147;
Matches 674; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
Qy       1686 GACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTG 1745
Db          1 GACAGCTGCTGGAAGCTGCAAAGGCTGGTGACGTTGTAACTGTAAAAAAACTCTGTACTG 60

Qy       1746 TTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAG 1805
Db         61 TTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCAGTCAACCCCACTCCACTTTGCAG 120

Qy       1806 CTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATG 1865
Db        121 CTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAACATGGAGCTGATGTTCATG 180

Qy       1866 CTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAG 1925
Db        181 CTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGCATGTTCTTATGGACACTATGAAG 240

Qy       1926 TTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTA 1985
Db        241 TTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCA 300

Qy       1986 CACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGC 2045
Db        301 CACCTTTACATGAAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGC 360

Qy       2046 ATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAG 2105
Db        361 ATGGTGCAGACCCTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAG 420

Qy       2106 ATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCA 2165
```

```
Db        421 ATGGAGATACAGATATTCAAGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCA 480

Qy       2166 AGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCG 2225
             ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||| |
Db        481 AGAAGGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTG 540

Qy       2226 ATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAG 2285
             |||||||||| | ||||| ||||||||||||||| |||||||||||||||||| || ||||
Db        541 ATACCCAAGGTCGGCATTCTACACCTTTACATTTGGCAGCTGGTTATAATAACTTGGAAG 600

Qy       2286 TTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTA 2345
             |||| |||||||| || ||||| |||||||||||||||||| |||||||||||||||| |
Db        601 TTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGACAAAGGAGGACTCA 660

Qy       2346 TTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGT 2405
             |||||||||||| |||||||| || || |||||||||||||||||| | || |||||||
Db        661 TTCCTTTACATAACGCAGCATCGTATGGACATGTAGATGTAGCAGCTTTGCTGATAAAGT 720

Qy       2406 ATAATGCATGTGTCAATGCCACG 2428
             | |  || ||||||||||||||||
Db        721 ACATGGCGTGTGTCAATGCCACG 743


RESULT 15
BX474686
LOCUS       BX474686                638 bp    mRNA    linear   EST 04-SEP-2003
DEFINITION  DKFZp686J06171_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686J06171 5', mRNA sequence.
ACCESSION   BX474686
VERSION     BX474686.1  GI:31668891
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp686J06171) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686J06171"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             16.4%;  Score 622.8;  DB 13;  Length 638;
  Best Local Similarity   99.5%;  Pred. No. 1.6e-144;
  Matches  635;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 60

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         61 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCAAG 120

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 180

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 240

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATAC 2085
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        241 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATAC 300

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGC 2145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGC 360

Qy       2146 AGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCC 2205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCC 420

Qy       2206 TGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC-AG 2264
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        421 TGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCNAG 480

Qy       2265 CTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATG 2324
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATG 540

Qy       2325 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATG 2384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATG 600

Qy       2385 TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAAT 2422
             |||||||||||||||||||||||||||||||||||||||
Db        601 TAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAAT 638


RESULT 16
CF533694
LOCUS       CF533694                724 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-FY0-cgs-i-17-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
            IMAGE:30362032 5', mRNA sequence.
ACCESSION   CF533694
VERSION     CF533694.1  GI:34585662
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
```

/db_xref="taxon:10090"
/clone="IMAGE:30362032"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 16.2%; Score 616.4; DB 14; Length 724;
Best Local Similarity 91.2%; Pred. No. 6.7e-143;
Matches 653; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```
Qy       1889 CCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACAT 1948
             ||||| |||||||||||||||||||||||||||||||||||||||||||||| || |||
Db          1 CCTTTACACAATGCATGTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCAT 60

Qy       1949 GGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCA 2008
             |||||||||||| ||||| ||||| || ||||| || |||||||||||||||||| ||||||
Db         61 GGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACCTTTACATGAAGCTGCAGCA 120

Qy       2009 AAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAA 2068
             |||||||||||||||| |||||||||| |||||||||||||||||||||||| || |||
Db        121 AAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGGTGCAGACCCTACAAAGAAA 180

Qy       2069 AACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGAT 2128
             ||||| ||||| ||||||||| ||||||||||||||||||||||||||||||||||||||
Db        181 AACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGAT 240

Qy       2129 CTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTG 2188
             || ||||| || |||||||||||||| ||||||||||||||||||||||||||||||||||
Db        241 CTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTG 300

Qy       2189 AAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACA 2248
             ||||| |||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        301 AAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACA 360

Qy       2249 CCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACAC 2308
             |||||||||| |||||||||||||||||| || |||||||| |||||||| || |||||
Db        361 CCTTTACATTTGGCAGCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACAT 420

Qy       2309 GGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCT 2368
             |||||||||||||||||||| |||||||||||||| |||||||||||| ||||||||||
Db        421 GGAGCTGATGTGAATGCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCG 480

Qy       2369 TACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACG 2428
             || || |||||||||||||||||| | || |||||||| ||||| ||||||||||||||||
Db        481 TATGGACATGTAGATGTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACG 540

Qy       2429 GACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGT 2488
             |||||||||||||||||||| || || |||||||||||||||||||| ||||||||||||
Db        541 GACAAATGGGCTTTCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGT 600

Qy       2489 GCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCT 2548
             ||||| ||| | |||||||||||||||| || ||||||| |||||| |||||||| ||||||
Db        601 GCTTTATTGTTGGCCCATGGAGCTGATCCTACTCTTANAAATCAAGAAGGACANACACCT 660

Qy       2549 TTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCC 2604
             |||||||||||||||||| |||||||||| || ||  ||||||||||||||||||||
```

Db 661 TTAGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCC 716

RESULT 17
CF727485
LOCUS CF727485 747 bp mRNA linear EST 09-OCT-2003
DEFINITION UI-M-HB0-ckk-b-13-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone IMAGE:30547140 5', mRNA sequence.
ACCESSION CF727485
VERSION CF727485.1 GI:37601653
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 747)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .747
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30547140"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 16.2%; Score 614.2; DB 14; Length 747;
Best Local Similarity 90.5%; Pred. No. 2.4e-142;
Matches 677; Conservative 0; Mismatches 69; Indels 2; Gaps 2;

```
Qy       1905 GTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATG 1964
             |||||||||||||| |||||||||||||||||||||||| || ||||||||||||| ||||
Db          1 GTTCTTATGGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATG 60

Qy       1965 TAGCTGATTTAT-GGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2023
             | ||||| || | |||| || |||||||||||||||||| ||||||||||||||||||||
Db         61 TCGCTGACTTGTNGGAAGTTCACACCTTTACATGAAGCTGCAGCAAAAGGAAAATATGAA 120

Qy       2024 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 2083
             || |||||||||| |||||||||||||||||||||||| || |||||||| ||||| |||
Db        121 ATATGCAAACTTCTACTCCAGCATGGTGCAGACCCTACAAAGAAAAACAGAGATGGTAAT 180
```

```
Qy       2084 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2143
             |||||| |||||||||||||||||||||||||||||||||||||||| ||||| || |||
Db        181 ACTCCTCTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTACTTAGAGGTGAT 240

Qy       2144 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2203
             |||||||||||| |||||||||||||||||||||||||||||||||||||| |||||||||
Db        241 GCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTCT 300

Qy       2204 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2263
             |||||||||||||||||| |||||||||| | ||||| |||||||||||||| |||
Db        301 CCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACACCTTTACATTTGGCA 360

Qy       2264 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2323
             |||||||||||||| || |||||||| |||||||| || ||||| ||||||||||||||
Db        361 GCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAAT 420

Qy       2324 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2383
             ||||| |||||||||||||| |||||||||||||||| |||||||| || || ||||||||
Db        421 GCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTAGAT 480

Qy       2384 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2443
             |||||||| | || |||||||| ||||| ||||||||||||||||||||||||||||||
Db        481 GTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTTC 540

Qy       2444 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2503
             ||||| || || |||||||||||||||||||| |||||||||||||||| ||| | |||
Db        541 ACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGCC 600

Qy       2504 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2563
             |||||||||| || |||||||||||||| |||||||| |||||||||||||||||||||
Db        601 CATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACANACACCTTTAGATTTAGTTTCA 660

Qy       2564 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2623
             || |||||||||| || ||  |||||||||||||||||||| ||||||||||  | ||
Db        661 GCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGC 720

Qy       2624 TACAAGCCTCAAGTGCTCAATGGTGTGA 2651
             |||||  |||||||||||| ||| ||||
Db        721 TACAA-ACTCAAGTGCTCAGTGGCGTGA 747
```

RESULT 18
BG875327

```
LOCUS       BG875327                 648 bp    mRNA    linear   EST 30-MAY-2001
DEFINITION  RC1-BT0623-120200-011-d04 BT0623 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG875327
VERSION     BG875327.1  GI:14252306
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC1-BT0623-120
            200-011-d04&t3=2000-02-12&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 4
            High quality sequence stop: 648.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="BT0623"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN
```

Query Match 15.9%; Score 605; DB 12; Length 648;
Best Local Similarity 98.7%; Pred. No. 4.7e-140;
Matches 620; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

```
Qy       1644 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1703
             |||| |||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          23 TCCTACAAGAGGGTATCTCATTAGG--ATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 80

Qy       1704 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1763
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 140

Qy       1764 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1823
             |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db         141 GAGACATTGAAGGGCGTCAGTCTACACCCNTTCATTTTGCAGCTGGGTATAACAGAGTGT 200

Qy       1824 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1883
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 260

Qy       1884 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1943
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         261 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 320

Qy       1944 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2003
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         321 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 380

Qy       2004 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2063
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         381 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 440

Qy       2064 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         441 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 500

Qy       2124 AAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCA 2183
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db         501 AAGATCTGCTTAGGGGAGATGCAGCTTAGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCA 560

Qy       2184 GAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATT 2243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         561 GAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATT 620

Qy       2244 CAACACCTTTACATTTAGCAGCTGGTTA 2271
             |||||||||||||| ||||||||||||
Db         621 CAACACCTTTACATATAGCAGCTGGTTA 648
```

RESULT 19
CD803519

```
LOCUS       CD803519                703 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  UI-M-GV0-chu-d-17-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
            IMAGE:30545656 5', mRNA sequence.
ACCESSION   CD803519
VERSION     CD803519.1  GI:32462345
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30545656"
                     /tissue_type="whole brain"
                     /dev_stage="1,5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GV0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction,ligated with
                     EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is CGAACTGAAT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match           15.7%;  Score 596;  DB 14;  Length 703;
  Best Local Similarity   90.6%;  Pred. No. 8.5e-138;
  Matches  635;  Conservative    0;  Mismatches   66;  Indels    0;  Gaps    0;

Qy       1930 AGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACC 1989
             ||||||||||||  || ||||||||||||||| ||||| |||||  || |||||  || ||||||
Db          2 AGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACC 61

Qy       1990 TTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGG 2049
             ||||||||||| |||||||||||||||||||||||||||| |||||||||| ||||||||||
Db         62 TTTACATGAAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGG 121

Qy       2050 TGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGG 2109
             ||||||||||| || |||||||| ||||| ||||||||| ||||||||||||||||||||
Db        122 TGCAGACCCTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGG 181

Qy       2110 AGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAA 2169
             |||||||||||||||||||| ||||| || |||||||||||||| ||||||||||||||
Db        182 AGATACAGATATTCAAGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAA 241

Qy       2170 GGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATAC 2229
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        242 GGGTTGTTTAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATAC 301

Qy       2230 CCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGC 2289
             |||||  | ||||| |||||||||||||| |||||||||||||||||| || ||||||||
Db        302 CCAAGGTCGGCATTCTACACCTTTACATTTGGCAGCTGGTTATAATAACTTGGAAGTTGC 361

Qy       2290 AGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCC 2349
              |||||||| || ||||| |||||||||||||||||||| |||||||||||||| |||||
Db        362 TGAGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGACAAAGGAGGACTCATTCC 421

Qy       2350 TTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAA 2409
             ||||||||| |||||||| || || |||||||||||||||||| | || |||||||| ||
Db        422 TTTACATAACGCAGCATCGTATGGACATGTAGATGTAGCAGCTTTGCTGATAAAGTACAA 481

Qy       2410 TGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAA 2469
             ||| |||||||||||||||||||||||||||||||||||| || || ||||||||||||||
Db        482 TGCGTGTGTCAATGCCACGGACAAATGGGCTTTCACACCGTTACATGAAGCAGCCCAAAA 541

Qy       2470 GGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAA 2529
             |||||| |||||||||||||||||| ||| | |||||||||||||| || ||||||||||
Db        542 GGGACGGACACAGCTTTGTGCTTTATTGTTGGCCCATGGAGCTGATCCTACTCTTAAAAA 601

Qy       2530 TCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGAC 2589
             ||| |||||||||||||||||||||||||||||||| |||||||||| || ||  ||||
Db        602 TCAAGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGAC 661

Qy       2590 AGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGC 2630
             ||||| |||||||| |||||||||| | || ||||| |
Db        662 AGCAGNCATGCCCCCCTCTGCTCTGCCTACGTGCTACAACC 702


RESULT 20
BE220005/c
LOCUS       BE220005                622 bp    mRNA    linear   EST 03-JUL-2000
DEFINITION  hv65h03.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3178325 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BE220005
VERSION     BE220005.1  GI:8907323
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 622)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1. .622
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3178325"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
```

modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 15.6%; Score 593.2; DB 10; Length 622;
Best Local Similarity 97.1%; Pred. No. 4.2e-137;
Matches 604; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy       3164 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3223
Db        622 GCAATCTTCACACGATACTATATTCTCAAGATTCAGCAGGTCTGTACAAGTACACTATCG 563

Qy       3224 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3283
Db        562 CAAAGATACACTCACGGCAGAAAAGCAGTTTCTGAAGAAAACCCCAACCATGCCAATGAA 503

Qy       3284 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3343
Db        502 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 443

Qy       3344 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3403
Db        442 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 383

Qy       3404 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3463
Db        382 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 323

Qy       3464 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3523
Db        322 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 263

Qy       3524 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3583
Db        262 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 203

Qy       3584 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643
Db        202 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 143

Qy       3644 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
Db        142 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 83

Qy       3704 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3763
Db          82 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 23

Qy       3764 CTCTACGTTTTACTCCTTTGCT 3785
Db          22 CTCTACGTTTTACTCCTTTGCT 1
```

RESULT 21
BE467629/c
LOCUS BE467629 616 bp mRNA linear EST 27-JUL-2000
DEFINITION hz67b11.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3213021 3' similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION BE467629
VERSION BE467629.1 GI:9513404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 616)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 402.
FEATURES Location/Qualifiers
source 1. .616
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3213021"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 15.4%; Score 584; DB 10; Length 616;
Best Local Similarity 96.8%; Pred. No. 8.4e-135;
Matches 596; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy       3170 TTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGA 3229
Db        616 TTCACAAGATCCAATATTCTCAAGATCATATAGGTTTGCAACCACGAACTATGGAATAGA 557

Qy       3230 TACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATG 3289
Db        556 TACACTCACCGGAGACAAGAAGTTTTTGAAGAAAACCCCACCCATGCCAATGACGGAATG 497

Qy       3290 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 3349
Db        496 CTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCAT 437

Qy       3350 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 3409
Db        436 GCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGC 377

Qy       3410 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 3469
Db        376 AATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCT 317

Qy       3470 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 3529
Db        316 TGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTG 257

Qy       3530 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 3589
Db        256 CAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGG 197

Qy       3590 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 3649
Db        196 CCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTAT 137

Qy       3650 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATA 3709
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 CCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATA 77

Qy       3710 GTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTAC 3769
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         76 GTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTAC 17

Qy       3770 GTTTTACTCCTTTGCT 3785
                ||||||||||||||||
Db         16 GTTTTACTCCTTTGCT 1


RESULT 22
BU466115
LOCUS       BU466115                793 bp    mRNA    linear   EST 30-NOV-2002
DEFINITION  603371794F1 CSEQRBN20 Gallus gallus cDNA clone ChEST279124 5', mRNA
            sequence.
ACCESSION   BU466115
VERSION     BU466115.1  GI:25959692
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 793)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .793
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer and broiler"
                     /db_xref="taxon:9031"
                     /clone="ChEST279124"
                     /sex="Male and female"
                     /tissue_type="Chondrocytes isolated from growth plate
                     cartilage"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQRBN20"
                     /note="Vector: pBluescript II KS(+); Site_1: EcoRI;
                     Site_2: NotI; This normalized library was constructed from
                     1 million independent clones. cDNA synthesis was initiated
                     using an oligo(dT) primer, using methylated C in the first
                     strand synthesis reaction. Following this first strand
                     reaction, double-stranded cDNA was blunted, ligated to
                     NotI adapters, digested with EcoRI, size-selected, and
                     cloned into the NotI and EcoRI compatible sites of a
                     custom modified MCS of the pBluescript (KS+) vector. The
                     library was normalized in 2 rounds using conditions
                     adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
                     Bonaldo et al., Genome Research 6 (1996): 791, except that
                     a significantly longer reannealing hybridization was
                     used."
ORIGIN

  Query Match           15.4%;  Score 583.8;  DB 13;  Length 793;
  Best Local Similarity   84.5%;  Pred. No. 9.8e-135;
  Matches  668;  Conservative    0;  Mismatches  122;  Indels    1;  Gaps    1;

Qy       1274 AATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATAT 1333
                || |||||||||||||||| ||||| |||||| |||| ||||||||||||||||| |||
Db          4 AACTTCAAGCATCCTCAAACGCATGAGACAGCACTGCACTGTGCTGCTGCATCTCCTTAT 63

Qy       1334 CCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAG 1393
                || || ||||| ||| |||||||| || |||| || |||||||| ||||| |||||||||
Db         64 CCTAAGAGAAAACAAGTATGTGAATTGCTGCTGAGGAAAGGAGCTAACATTAATGAAAAG 123

Qy       1394 ACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTT 1453
                || ||||| |||||||||||||||| |||||||| || ||||||||||||||||||||
Db        124 ACAAAAGACTTCTTGACTCCTCTGCATGTGGCATCAGAAAAAGCTCATAATGATGTTGTT 183

Qy       1454 GAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCT 1513
                |||||||| |||||||||||||||||||||||||||| ||||||| ||||| ||||| |||
Db        184 GAAGTAGTTGTGAAACATGAAGCAAAGGTTAATGCATTGGATAACCTTGGCCAGACCTCT 243

Qy       1514 CTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGG 1573
                || |||||||| ||| |||||||||| || || ||| | | | |||| | |||
Db        244 CTTCACAGAGCGGCACATTGTGGTCACCTCCAGACTTGCAGGTTGTTGTTGAGTTCTGGA 303

Qy       1574 TGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAAT 1633
                |||||||| |||| | || | |||||||||| || ||||| ||||| ||||||| |
Db        304 TGTGATCCTTCCATTGTGTCTTTGCAGGGCTTTACAGCCTTACAAATGGGGAATGAAAGT 363

Qy       1634 GTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTG 1693
                || || ||||| || ||||| |||||| | |||||||||||| ||||| |||||||||
Db        364 GTGCAACAACTACTGCAAGAAGGTATCCCTCTAGGTAATTCAGATGCAGATAGACAATTG 423

Qy       1694 CTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGT 1753
                |||| ||||| |||||||| ||||| || ||||||||||||||||||| ||||| |||
Db        424 TTGGAGGCTGCGAAGGCTGGCGATGTGGATACTGTAAAAAAACTGTGTACGGTTCAAAGT 483

Qy       1754 GTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTAT 1813
                || |||||||||| |||||||||||||||||||||||||||||||||||||||| ||
Db        484 GTGAACTGCAGAGATATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGATAC 543

Qy       1814 AACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGAT 1873
                || || ||||| || || |||||||| || |||||||||||||||||||||||| ||||||
Db        544 AATAGGGTGTCTGTTGTTGAATATCTTCTTCAGCATGGAGCTGATGTGCATGCAAAAGAT 603

Qy       1874 AAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAA 1933
                || |||||| ||||| ||||| || || || || || |||||| || ||||| ||||| ||
Db        604 AAGGGAGGACTTGTTCCTTTACATAACGCTTGCTCATATGGTCACTATGAGGTTGCGGAG 663

Qy       1934 CTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTA 1993
                || |||||||||||||| |||||| || || |||||||| || ||||| |||||| || ||
Db        664 CTGCTTGTTAAACATGGTGCAGTTGTCAACGTAGCTGACTTGTGGAAGTTTACTCCCTTG 723

Qy       1994 CATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCA 2053
                |||||||| |||||||||||||||||||| |||||| |||| ||   | |||||||| ||
Db        724 CATGAAGCTGCAGCAAAAGGAAAATATGAGATTTGC-AACTCCTATTGCAGCATGGAGCT 782

Qy       2054 GACCCTACCAA 2064
                ||||| |  ||
Db        783 GACCCCAAAAA 793


RESULT 23
CF739820
LOCUS       CF739820                689 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-HD0-clf-m-21-0-UI.r1 NIH_BMAP_HD0 Mus musculus cDNA clone
            IMAGE:30611156 5', mRNA sequence.
ACCESSION   CF739820
VERSION     CF739820.1  GI:37636158
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .689
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30611156"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HD0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I;  The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH)."
ORIGIN

  Query Match             15.2%;  Score 578.8;  DB 14;  Length 689;
  Best Local Similarity   90.0%;  Pred. No. 1.7e-133;
  Matches  619;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy       2898 AGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCA 2957
              |||| || || |||||  ||||||| ||||| |||||||| ||||| ||||||||||||
Db          1 AGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCA 60

Qy       2958 ATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC 3017
              |||| |||||||||  | |||||  || ||||||||||| || || || ||||| |||||||
Db         61 ATGCCTATGGACATCGACACAAGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAAC 120

Qy       3018 AAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATC 3077
              |||||||||| |||||||||||  | ||||  |||||||||||||||||||||| |||||||
Db        121 AAGGTCTTAATCCATATTTAACTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATC 180

Qy       3078 TGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAG 3137
              |||||||||||||||||  |||||||||||||  |||||||||| |||||  ||  ||  |
Db        181 TGTCTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGG 240

Qy       3138 AGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTC 3197
              |||||||||||| |||||  |||||  ||  ||||||||||||||||  ||||||||||||
Db        241 AGCACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTC 300

Qy       3198 AGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTG 3257
              |||||||||||||||||  | ||||||||||  ||  ||||||||||||||||||||||
Db        301 AGAAGGTTTGTAACAAGAAATTGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTG 360

Qy       3258 AAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATG 3317
              |||||||||||||||||  || |||  | |||  |||||||||||||||||||||||||||
Db        361 AAGAAAACCACAACCATGCAAACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATG 420

Qy       3318 CAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTG 3377
              |  |||||||  || |||||||||||||||||||  ||||||||||  ||||||||||||
Db        421 CGATTATCCATAAGGGCTTTGATGAAAGGCATGCATACATAGGTGGCATGTTTGGAGCTG 480

Qy       3378 GCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAG 3437
              | |||||||||||| |||||||||||||||||||||||||||| |||||||||||||| |
Db        481 GAATTTATTTTGCTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGAATTGGAGGTG 540

Qy       3438 GTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCT 3497
              | || || |||||| |||||||||||||||| |||||||||| |||||||||||||| |
Db        541 GCACCGGATGTCCAATTCACAAAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCTGT 600

Qy       3498 TTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATT 3557
              |||| || |||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        601 TTTGTCGAGTAACCTTGGGAAAGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATT 660

Qy       3558 CTCCTCCAGGTCATCACTCAGTCACTGG 3585
              |||||| || |||||||| ||||||||
Db        661 CTCCTCNCGGNCATCACTCGGTCACTGG 688


RESULT 24
BU427328
LOCUS       BU427328                796 bp    mRNA    linear   EST 29-NOV-2002
DEFINITION  603234467F1 CSEQRBN09 Gallus gallus cDNA clone ChEST235e17 5', mRNA
            sequence.
ACCESSION   BU427328
VERSION     BU427328.1  GI:25920004
KEYWORDS    EST.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
  TITLE     A Comprehensive Collection of Chicken cDNAs
  JOURNAL   Curr. Biol. 12 (22), 1965-1969 (2002)
  MEDLINE   22335534
   PUBMED   12445392
COMMENT     Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer and broiler"
                     /db_xref="taxon:9031"
                     /clone="ChEST235e17"
                     /sex="Male and female"
                     /tissue_type="Chondrocytes isolated from growth plate
                     cartilage"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="CSEQRBN09"
                     /note="Vector: pBluescript II KS(+); Site_1: EcoRI;
                     Site_2: Not1; This normalized library was constructed from
                     1 million independent clones. cDNA synthesis was initiated
                     using an oligo(dT) primer, using methylated C in the first
                     strand synthesis reaction. Following this first strand
                     reaction, double-stranded cDNA was blunted, ligated to
                     NotI adapters, digested with EcoRI, size-selected, and
                     cloned into the NotI and EcoRI compatible sites of a
                     custom modified MCS of the pBluescript (KS+) vector. The
                     library was normalized in 2 rounds using conditions
                     adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
                     Bonaldo et al., Genome Research 6 (1996): 791, except that
                     a significantly longer reannealing hybridization was
                     used."
```

ORIGIN

Query Match 15.1%; Score 574; DB 13; Length 796;
Best Local Similarity 83.5%; Pred. No. 2.8e-132;
Matches 663; Conservative 0; Mismatches 130; Indels 1; Gaps 1;

```
Qy       1274 AATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATAT 1333
             || |||||||||||||||||| ||||| |||||| |||| ||||||||||||||||||| |||
Db          4 AACTTCAAGCATCCTCAAACGCATGAGACAGCACTGCACTGTGCTGCTGCATCTCCTTAT 63

Qy       1334 CCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAG 1393
             || || ||||| ||| |||||||| || |||| || |||||||| |||| |||||||||
Db         64 CCTAAGAGAAAACAAGTATGTGAATTGCTGCTGAGGAAAGGAGCTAACATTAATGAAAAG 123

Qy       1394 ACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTT 1453
             || |||||| |||||||||||||||| |||||||| || ||||||||||||||||||||
Db        124 ACAAAAGACTTCTTGACTCCTCTGCATGTGGCATCAGAAAAAGCTCATAATGATGTTGTT 183

Qy       1454 GAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCT 1513
             |||||||| |||||||||||||||||||||||||| |||||| |||| ||||| |||
Db        184 GAAGTAGTTGTGAAACATGAAGCAAAGGTTAATGCATTGGATAACCTTGGCCAGACCTCT 243

Qy       1514 CTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGG 1573
             || |||||||| ||| |||||||||| || || || ||| |  |  | |||| | |||
Db        244 CTTCACAGAGCGGCACATTGTGGTCACCTCCAGACTTGCAGGTTGTTGTTGAGTTCTGGA 303

Qy       1574 TGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAAT 1633
             |||||||||  |||| | ||  | |||||||||| || ||||| ||||| |||||||
Db        304 TGTGATCCTTCCATTGTGTCTTTGCAGGGCTTTACAGCCTTACAAATGGGGAATGAAAGT 363

Qy       1634 GTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTG 1693
             || || ||||| || ||||| |||||| |  |||||||||||| ||||| |||||||||
Db        364 GTGCAACAACTACTGCAAGAAGGTATCCCTCTAGGTAATTCAGATGCAGATAGACAATTG 423

Qy       1694 CTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGT 1753
              |||| ||||| |||||||| ||||| || |||||||||||||||||||| ||||| |||
Db        424 TTGGAGGCTGCGAAGGCTGGCGATGTGGATACTGTAAAAAAACTGTGTACGGTTCAAAGT 483

Qy       1754 GTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTAT 1813
             || |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||
Db        484 GTGAACTGCAGAGATATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGATAC 543

Qy       1814 AACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGAT 1873
             || || ||||| || || |||||||| || ||||||||||||||||||||| ||||||
Db        544 AATAGGGTGTCTGTTGTTGAATATCTTCTTCAGCATGGAGCTGATGTGCATGCAAAAGAT 603

Qy       1874 AAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAA 1933
             || |||||| |||| |||| || || || || || ||||| || ||||| ||||| ||
Db        604 AAGGGAGGACTTGTTCCTTTACATAACGCTTGCTCATATGGTCACTATGAGGTTGCGGAG 663

Qy       1934 CTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTA 1993
             || |||||||||||||| |||| || |||||||| || ||||| ||||| || ||
Db        664 CTGCTTGTTAAACATGGTGCAGTTGTCAACGTAGCTGACTTGTGGAAGTTTACTCCCTTG 723

Qy       1994 CATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCA 2053
             |||||||| |||| |||||||||||||| |||||||    || | ||||||||
Db        724 CATGAAGCTGCAGC-AAAGGAAAATATGAGATTTGCAACTCCCTATTGCAGCATGGAAGC 782

Qy       2054 GACCCTACCAAAAA 2067
                 ||  | |||||
Db        783 TGACCCCCAAAAAA 796
```

RESULT 25
CF735419
LOCUS CF735419 742 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-HB0-cke-a-13-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone IMAGE:30549036 5', mRNA sequence.
ACCESSION CF735419
VERSION CF735419.1 GI:37631755
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 742)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .742
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30549036"
/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 15.1%; Score 572.8; DB 14; Length 742;
Best Local Similarity 86.5%; Pred. No. 5.5e-132;
Matches 642; Conservative 0; Mismatches 99; Indels 1; Gaps 1;

```
Qy       2274 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2333
             |||| || |||||||| |||||||| || ||||| |||||||||||||||||||| ||||
Db          1 ATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGACA 60

Qy       2334 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2393
             |||||||||| |||||||||||| |||||||| || || ||||||||||||||||||
Db         61 AAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTAGATGTAGCAGCTT 120

Qy       2394 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2453
             | || |||||||| ||||| |||||||||||||||||||||||||||||||||| || |
Db        121 TGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTTCACACCGTTAC 180

Qy       2454 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2513
             | |||||||||||||||||| |||||||||||||| ||| | ||||||||||||||
Db        181 ATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGCCCATGGAGCTG 240

Qy       2514 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2573
             | || |||||||||||||| |||||||||||||||||||||||||||||||| |||||||
Db        241 ATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 300

Qy       2574 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2633
             |||| || ||  |||||||||||||||||| |||||||||| | || ||||| ||||
Db        301 TCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGCTACAAACCTC 360

Qy       2634 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2693
```

```
Db        361 AAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCATTGCAGATGCTCTGTCTTCAGGTCCGT 420

Qy       2694 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2753

Db        421 CCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTGGCAGCTTCTCGG 480

Qy       2754 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2813

Db        481 AACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTTTGCANAGAAAAG 540

Qy       2814 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2873

Db        541 AGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACT 600

Qy       2874 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2933

Db        601 TGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACA 660

Qy       2934 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 2993

Db        661 AGGAACTGAAAGAGATTGG-ATCAATGCCTATGGACATCGATACANGCTGATTAAAGGAG 719

Qy       2994 TCGAGAGACTTATCTCCGGACA 3015

Db        720 TTGAAAGGCTGATCTCTGGACA 741
```

RESULT 26
CB201923
LOCUS CB201923 913 bp mRNA linear EST 05-FEB-2003
DEFINITION AGENCOURT_11289292 NIH_MGC_135 Mus musculus cDNA clone IMAGE:30140366 5', mRNA sequence.
ACCESSION CB201923
VERSION CB201923.1 GI:28236420
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 913)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0046 row: e column: 15
High quality sequence stop: 676.
FEATURES Location/Qualifiers
source 1. .913
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30140366"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/note="Vector: pCMVSport6.1; Site_1: EcoRV; Site_2: NotI; Normalized full-length enriched library from pooled mouse embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5, and 15.5 (size selected for the 0.5-1 kb fragments) Cloned directionally, priming method: Oligo-dT. cDNA enrichment: >1k bp, Average insert size 1.6k bp. Normalization (Cot value): 7.5 kb. Priming sequence: 5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)3' Tissue contributed by, David Rowe. Library constructed by ResGen, Invitrogen Corp."

ORIGIN

Query Match 15.1%; Score 572.6; DB 14; Length 913;
Best Local Similarity 91.8%; Pred. No. 6.4e-132;
Matches 605; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```
Qy       3049 CTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGT 3108

Db          2 CTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGT 61

Qy       3109 GGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAAT 3168

Db         62 GGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGT 121

Qy       3169 CTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAG 3228

Db        122 CTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTGGGAAAG 181

Qy       3229 ATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAAT 3288

Db        182 ATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAACGAAAGGAT 241

Qy       3289 GCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCA 3348

Db        242 GTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTTTGATGAAAGGCA 301

Qy       3349 TGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAG 3408

Db        302 TGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTCCAAAAG 361

Qy       3409 CAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATC 3468

Db        362 CAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTCCAATTCACAAAGACAGATC 421

Qy       3469 TTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCT 3528

Db        422 GTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAACCTTGGGAAAGTCTTTCTT 481

Qy       3529 GCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAG 3588

Db        482 GCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCCATCACTCGGTCACTGGCCG 541

Qy       3589 GCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTA 3648

Db        542 GCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTA 601

Qy       3649 TCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAA 3707

Db        602 TCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAGGTATGGTTGATGGCTGAA 660
```

RESULT 27
BE910901
LOCUS BE910901 956 bp mRNA linear EST 29-SEP-2000
DEFINITION 601661978F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3962251 5', mRNA sequence.
ACCESSION BE910901
VERSION BE910901.1 GI:10407885
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 956)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

```
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM9128  row: k  column: 20
          High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .956
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3962251"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match             15.0%;  Score 570.4;  DB 10;  Length 956;
  Best Local Similarity   85.2%;  Pred. No. 2.3e-131;
  Matches  673;  Conservative    0;  Mismatches  111;  Indels    6;  Gaps    3;

Qy       2828 GATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTT 2887
              ||||||||  ||||||||  |||  |||||||  |||||||||||||||||  | ||||||||||||
Db          1 GATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACTTGATGGATATATTT 60

Qy       2888 GAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAG 2947
              |||||||||||||| || || |||||  |||||||  ||||  ||||||||  |||||  |||
Db         61 GAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAAGGAACTGAAAGAG 120

Qy       2948 ATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATC 3007
              ||||||||||||||  |||||||||  |  |||||  ||  ||  ||||||||||||  |||  |||
Db        121 ATTGGAATCAATGCCTATGGACATCGACACAAGCTGATTAAAGGAGTTGAAAGGCTGATC 180

Qy       3008 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 3067
              ||  ||||||||||||||||||  ||||||||||||  |  ||||  ||||||||||||||||||
Db        181 TCTGGACAACAAGGTCTTAATCCATATTTAACTCTAAACAACTCTGGTAGTGGAACAATT 240

Qy       3068 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 3127
              ||  ||||||||||||||||||||||||||||  ||||||||||||||  ||||||||||  |||
Db        241 CTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGT 300

Qy       3128 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 3187
              ||  ||  ||  ||||||||||||||  |||||  |||||  ||  ||||||||||||||||  |||
Db        301 ACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATT 360

Qy       3188 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 3247
              ||||||||||||||||||||||||||||||  |  ||||||||||  ||  ||||||||||||
Db        361 CTCAAGATTCAGAAGGTTTGTAACAAGAAATTGTGGGAAAGATATACACACCGGAGAAAA 420

Qy       3248 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 3307
              ||||||||||||||||||||||||||||||  ||  |||  |  |||  ||||||||||||||||||
Db        421 GAAGTTTCTGAAGAAAACCACAACCATGCAAACGAAAGGATGTTATTTCATGGGTCTCCT 480

Qy       3308 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATG 3367
              ||||||||||  ||||||||  ||  ||||||||||||||||||||  ||||||||||  |||
Db        481 TTTGTGAATGCGATTATCCATAAGGGCTTTGATGAAAGGCATGCATACATAGGTGGCATG 540

Qy       3368 TTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGA 3427
              ||||||||||  ||||||||||||||  ||||||||||||||||||||||||||  ||||||
Db        541 TTTGGAGCTGGAATTTATTTTGCTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGA 600

Qy       3428 ATTGGAGGAGG---TACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCAC 3484
              |||  |  |  |  |  ||  ||||||  ||||||||||||||||  ||||||  ||||  |||
Db        601 ACTTGGAGGTGGCACCCGGGATGTCCAATTCACAAAGACAGATCGTGTTACCTTTGTCAC 660

Qy       3485 AGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATG 3544
              ||||||||||  |||  |  ||  ||||||||  |||||||||  |||  |||||  |||  ||
Db        661 AGGCAGCTGCTGTTTCGTCGAGT-ACCTTGGGCAAGTCTTTCTTGCCGTTCAACGCATTG 719

Qy       3545 AAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGG--TAGGCCCAGTGTAAATG 3602
              |  ||||||    |||||  ||  |       |    |  |||||||    ||||      |||||
Db        720 ACAATGGACATTTCTCTCCCGGCCATCAATCGGGCCACTGGCCCGGGCCAGTGTGAAATG 779

Qy       3603 GCCTAGCATT 3612
              ||||||  |||
Db        780 GCCTAGGATT 789


RESULT 28
BB339554
LOCUS       BB339554                 702 bp    mRNA    linear   EST 24-OCT-2001
DEFINITION  BB339554 RIKEN full-length enriched, 10 days neonate cerebellum Mus
            musculus cDNA clone B930018O21 3', mRNA sequence.
ACCESSION   BB339554
VERSION     BB339554.2  GI:16404413
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 702)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 11, 2000 this sequence version replaced gi:9048317.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
            Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
            and Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
            Sugahara,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
            Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .702
```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="B930018O21"
/sex="mixed"
/tissue_type="cerebellum"
/dev_stage="10 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days neonate cerebellum"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

ORIGIN

Query Match 15.0%; Score 568; DB 10; Length 702;
Best Local Similarity 89.7%; Pred. No. 8.6e-131;
Matches 621; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

```
Qy       1313 TGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAA 1372
             |||||||||||||||| |||||||||||||||||| |||||||||| |||| |||||
Db         12 TGTGCTGCTGCATCT-CATATCCCAAAAGAAAGCAGATATGTGAACTGCTGCTCAGAAAG 70

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
             |||||||||| ||| ||||||||||||||||||||||||||||||||||||||| ||||||
Db         71 GGAGCAAACACCAACGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCTTCTGAG 130

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
             || ||||| |||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        131 AACGCTCACAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTTTG 190

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
             |||| |||||| ||||| ||  |||||||||||||| | |||||||| || || ||||||
Db        191 GATAGTCTTGGACAGACGTCATTACACAGAGCTGCACACTGTGGTCACCTGCAGACCTGC 250

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
             ||||| |||||||||||||||||||||| ||||| |||||||||||||| || || ||
Db        251 CGCCTGCTCCTGAGCTATGGGTGTGATCCAAACATCATATCCCTTCAGGGTTTCACCGCC 310

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
             || || |||||||||||||||| ||||| || ||||||||||    |||| ||||| |
Db        311 TTGCAAATGGGAAATGAAAATGTGCAGCAGCTGCTCCAAGAGGGCGCCTCACTAGGTCAC 370

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAA 1732
             |||||||| |||||||| |||||||||||||||||||||| || || ||||||||||||
Db        371 TCAGAGGCGGACAGACAGCTGCTGGAAGCTGCAAAGGCTGGTGACGTTGAAACTGTAAAA 430

Qy       1733 AAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCA 1792
             ||||| |||||||||||||||||||||||||||||||||||||||||| || ||||| || |||
Db        431 AAACTCTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGACGCCAGTCAACCCCA 490

Qy       1793 CTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGA 1852
             || || |||||||||||||| |||||||||| ||||| |||||||||| || ||||||
Db        491 CTCCACTTTGCAGCTGGGTACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAACATGGA 550

Qy       1853 GCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTAT 1912
             |||||||| |||||||||| |||||||| |||||||||| ||||||||||||||||||
```

```
Db        551 GCTGATGTTCATGCTAAAGACAAAGGAGGTCTTGTACCTTTACACAATGCATGTTCTTAT 610

Qy       1913 GGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGAT 1972
             ||||| |||||||||||||||||||||| || |||||||||||||| ||||| ||||
Db        611 GGACACTATGAAGTTGCAGAACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGAC 670

Qy       1973 TTATGGAAATTTACACCTTTACATGAAGCAGC 2004
             || ||||| || |||||||||||||||||| ||
Db        671 TTGTGGAAGTTCACACCTTTACATGAAGCTGC 702
```

RESULT 29
CF538288
LOCUS CF538288 725 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-GH0-chd-p-24-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone IMAGE:30535583 5', mRNA sequence.
ACCESSION CF538288
VERSION CF538288.1 GI:34590270
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 725)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousef1.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .725
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30535583"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 14.9%; Score 564.2; DB 14; Length 725;
Best Local Similarity 86.9%; Pred. No. 7.8e-130;
Matches 631; Conservative 0; Mismatches 94; Indels 1; Gaps 1;

```
Qy       2439 CTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGC 2498
             |||||||||| || || |||||||||||||||||| |||||||||||||||| |||
Db          1 CTTTCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGT 60
```

```
Qy       2499 TAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAG 2558
Db         61 TGGCCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAG 120

Qy       2559 TTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCT 2618
Db        121 TTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTA 180

Qy       2619 CTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTC 2678
Db        181 CGTGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTC 240

Qy       2679 TCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTAT 2738
Db        241 TGTCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTAT 300

Qy       2739 CTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCA 2798
Db        301 CTGGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTG 360

Qy       2799 GTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATC 2858
Db        361 GTTTGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACC 420

Qy       2859 TTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAG 2918
Db        421 TTGGACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAG 480

Qy       2919 TTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACA 2978
Db        481 TTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACA 540

Qy       2979 AACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAA 3038
Db        541 AGCTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAA 600

Qy       3039 CTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGT 3098
Db        601 CTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGANTAAGA-T 659

Qy       3099 TTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATG 3158
Db        660 TTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACG 719

Qy       3159 CAGGTG 3164
Db        720 CAGGCG 725
```

RESULT 30
BE222311/c

```
LOCUS       BE222311                585 bp    mRNA    linear   EST 03-JUL-2000
DEFINITION  hu10d01.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3166177 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION   BE222311
VERSION     BE222311.1  GI:8909629
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 585)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 476.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3166177"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN
```

Query Match 14.8%; Score 562.8; DB 10; Length 585;
Best Local Similarity 98.5%; Pred. No. 1.7e-129;
Matches 576; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

```
Qy       3203 GTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAA 3262
Db        585 GTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAA 526

Qy       3263 AACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATT 3322
Db        525 AACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATT 466

Qy       3323 ATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATT 3382
Db        465 ATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATT 406

Qy       3383 TATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA--GGTA 3440
Db        405 TATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGNAGGAAGGNGN 346

Qy       3441 CTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTT 3500
Db        345 NNCGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTT 286

Qy       3501 GCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTC 3560
Db        285 GCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTC 226

Qy       3561 CTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAAT 3620
Db        225 CTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAAT 166

Qy       3621 ATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGA 3680
Db        165 ATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGA 106

Qy       3681 GGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCT 3740
Db        105 GGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCT 46

Qy       3741 AAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 3785
Db         45 AAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 1
```

RESULT 31

```
CF728039
LOCUS       CF728039                681 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  UI-M-HB0-ckm-c-21-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
            IMAGE:30548708 5', mRNA sequence.
ACCESSION   CF728039
VERSION     CF728039.1  GI:37602207
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 681)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30548708"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HB0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH)."
ORIGIN

  Query Match           14.8%;  Score 561.2;  DB 14;  Length 681;
  Best Local Similarity   89.9%;  Pred. No. 4.3e-129;
  Matches  613;  Conservative    0;  Mismatches   68;  Indels    1;  Gaps    1;

Qy       1578 ATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTAC 1637
             |||| ||||| |||||||||||||| || || || || || ||||||||||||||||||| |
Db          1 ATCCAAACATCATATCCCTTCAGGGTTTCACCGCCTTGCAAATGGGAAATGAAAATGTGC 60

Qy       1638 AGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGG 1697
             |||| || |||||||||||    |||| ||||| | |||||||| |||||||  ||||||
Db         61 AGCAGCTGCTCCAAGAGGGCGCCTCACTAGGTCACTCAGAGGCGGACAGACAGCTGCTGG 120

Qy       1698 AAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCA 1757
             |||||||||||||||| || || |||||||||||||||| |||||||||||||||||||||
Db        121 AAGCTGCAAAGGCTGGTGACGTTGAAACTGTAAAAAAACTCTGTACTGTTCAGAGTGTCA 180

Qy       1758 ACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACA 1817
             |||||||||||||||||| || ||||| || ||||| || |||||||||||||| ||||
Db        181 ACTGCAGAGACATTGAAGGACGCCAGTCAACCCCACTCCACTTTGCAGCTGGGTACAACA 240

Qy       1818 GAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAG 1877
             ||||||| ||||| |||||||||||| || |||||||||||||| |||||||||| ||||
Db        241 GAGTGTCTGTGGTAGAATATCTGCTGCAACATGGAGCTGATGTTCATGCTAAAGACAAAG 300

Qy       1878 GAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTC 1937
             |||| |||||||||| |||||||||||||||||||||||||||||| ||||| ||||||||||
Db        301 GAGGTCTTGTACCTTTACACAATGCATGTTCTTATGGACACTATGATGTTGCAGAACTTC 360

Qy       1938 TTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATG 1997
             |||| || |||||||||||||| ||||| ||||| || ||||| || ||||||||||||||
Db        361 TTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACCTTTACATG 420

Qy       1998 AAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACC 2057
             |||| |||||||||||||||||||||||| |||||||||| ||||||||||||||||||||
Db        421 AAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGGTGCAGACC 480

Qy       2058 CTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAG 2117
             |||| || |||||||| ||||| |||||||| |||||||||||||||||||||||||||||
Db        481 CTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGGAGATACAG 540

Qy       2118 ATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTT 2177
             ||||||| ||||| ||||| || |||||||||||||| |||||||||||||||||||||||
Db        541 ATATTCACGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAAGGGTTGTT 600

Qy       2178 TAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCA 2237
             |||||||||||||||| |||||||||||||||||||||||||||||||| |||| ||||||
Db        601 TAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATA-CCAAGGTC 659

Qy       2238 GACATTCAACACCTTTACATTT 2259
             | ||||| ||||||||||||||
Db        660 GGCATTCTACACCTTTACATTT 681


RESULT 32
CF745978
LOCUS       CF745978                669 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-GV0-clx-d-21-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
            IMAGE:30618620 5', mRNA sequence.
ACCESSION   CF745978
VERSION     CF745978.1  GI:37642321
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30618620"
                     /tissue_type="whole brain"
                     /dev_stage="1,5, and 15 days newborn"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GV0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
```

Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction,ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 14.7%; Score 557.8; DB 14; Length 669;
Best Local Similarity 90.4%; Pred. No. 3.1e-128;
Matches 606; Conservative 0; Mismatches 63; Indels 1; Gaps 1;

```
Qy       970 TCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTT 1029
            |||||||||||||||||||||||| |||||||||| |||||||||| |||||||||||| |
Db         1 TCATTATGAAGTAACTGAACTTCTGGTCAAGCACGGTGCCTGTGTGAATGCAATGGACCT 60

Qy      1030 GTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCT 1089
            |||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        61 GTGGCAGTTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGATTGAAGTATGTTCTCT 120

Qy      1090 TCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGA 1149
               ||| ||||||| || ||||| ||||| ||||| || ||||||||||||||||||||||
Db       121 CCTCCTAAGTTACGGGGCAGATCCAACCCTGCTGAACTGTCACAATAAAAGTGCTATAGA 180

Qy      1150 CTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTT 1209
            | |||| ||||| |||||  | || ||||||||| | ||||||||||||||||| || ||
Db       181 CCTGGCCCCCACGGCACAGCTGAAGGAAAGATTATCGTATGAATTTAAAGGCCATTCATT 240

Qy      1210 GCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAAT 1269
            |||||||||||||||||||||||||||||||  |||||||||||||| ||||||||||||
Db       241 GCTGCAAGCTGCACGAGAAGCTGATGTTACAAGAATCAAAAAACATCTTTCTCTGGAAAT 300

Qy      1270 GGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCC 1329
            |||||||||||| |||||||| |||||||||||||| |||||||||||||||||||||||
Db       301 GGTGAATTTCAAACATCCTCAGACACATGAAACAGCACTGCATTGTGCTGCTGCATCTCC 360

Qy      1330 ATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGA 1389
            |||||||||||||||||||||||||||||||| |||| |||| |||||||||| ||| ||
Db       361 ATATCCCAAAAGAAAGCAGATATGTGAACTGCTGCTCAGAAAGGGAGCAAACACCAACGA 420

Qy      1390 AAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGT 1449
            |||||||||||||||||||||||||||||||||||||||| ||||||| ||||| ||||||||
Db       421 AAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCTTCTGAGAACGCTCACAATGATGT 480

Qy      1450 TGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGAC 1509
            |||||||||||||||||||||||||||||||||||||||||||||| |||||| |||||||
Db       481 TGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTTTGGATAGTCTTGGACAGAC 540

Qy      1510 TTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTA 1569
               ||||||||||||||| |||||||| || || |||||||||| ||||||||||||||||
Db       541 GTCATTACACAGAGCTGCACACTGTGGTCACCTGCAGACCTGCCGCCTGCTCCTGAGCTA 600

Qy      1570 TGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGA 1629
            |||||||||| |  |||| |||||||||||||| || || || || || |||||| ||||
Db       601 TGGGTGTGAT-CANACATCATATCCCTTCAGGGTTTCACCGCCTTGCATATGGGACATGA 659

Qy      1630 AAATGTACAG 1639
            |||||| |||
Db       660 AAATGTGCAG 669
```

RESULT 33
BE504175/c
LOCUS BE504175 595 bp mRNA linear EST 04-AUG-2000
DEFINITION hv86c11.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3180308 3' similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION BE504175
VERSION BE504175.1 GI:9706583
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 409.
FEATURES Location/Qualifiers
source 1. .595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3180308"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 14.7%; Score 557.4; DB 10; Length 595;
Best Local Similarity 99.3%; Pred. No. 3.8e-128;
Matches 591; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

```
Qy      3194 ATTCAGAAGGTTTGTAACAAG-AAACTATGGGAAAGATACACTCACCGG-AGAAAAGAAG 3251
            |||||||||||||||||||||| |||||||||||||||||||||||||| | ||||||||||
Db       595 ATTCAGAAGGTTTGTAACAAGAAAACTATGGGAAAGATACACTCACCCGAAGAAAAGAAG 536

Qy      3252 TTTCTGAAGAAAACCAC-AACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 3310
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       535 TTTCTGAAGAAAACCACAAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTT 476

Qy      3311 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 3370
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       475 GTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTT 416

Qy      3371 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 3430
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       415 GGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATT 356

Qy      3431 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 3490
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       355 GGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAG 296

Qy      3491 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 3550
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       295 CTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATG 236
```

```
Qy       3551 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 3610
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCA 176

Qy       3611 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 3670
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 TTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTAC 116

Qy       3671 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 3730
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 CAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTC 56

Qy       3731 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 3785
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 CACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCT 1
```

RESULT 34
CF530741
LOCUS CF530741 640 bp mRNA linear EST 12-SEP-2003
DEFINITION UI-M-FY0-cgm-g-09-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone IMAGE:30354680 5', mRNA sequence.
ACCESSION CF530741
VERSION CF530741.1 GI:34582705
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 640)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.

FEATURES Location/Qualifiers
source 1. .640
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30354680"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 14.5%; Score 552; DB 14; Length 640;
Best Local Similarity 91.4%; Pred. No. 8.6e-127;
Matches 585; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

```
Qy       1692 TGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGA 1751
             |||||||||||||||||||| || || |||||||||||||||| ||||||||||||||
Db          1 TGCTGGAAGCTGCAAAGGCTGGTGACGTTGAAACTGTAAAAAAACTCTGTACTGTTCAGA 60

Qy       1752 GTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGT 1811
             |||||||||||||||||||||||||| || ||||| || ||||| || ||||||||||||
Db         61 GTGTCAACTGCAGAGACATTGAAGGACGCCAGTCAACCCCACTCCACTTTGCAGCTGGGT 120

Qy       1812 ATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAG 1871
             | |||||||||| |||| |||||||||| || |||||||||||||| ||||||||||
Db        121 ACAACAGAGTGTCTGTGGTAGAATATCTGCTGCAACATGGAGCTGATGTTCATGCTAAAG 180

Qy       1872 ATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAG 1931
             | |||||||| |||||||||| ||||||||||||||||||||||||| ||||||||||||
Db        181 ACAAAGGAGGTCTTGTACCTTTACACAATGCATGTTCTTATGGACACTATGAAGTTGCAG 240

Qy       1932 AACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTT 1991
             |||||||||| || |||||||||||||| |||| ||||| || ||||| || |||||||
Db        241 AACTTCTTGTCAAGCATGGAGCAGTAGTCAATGTCGCTGACTTGTGGAAGTTCACACCTT 300

Qy       1992 TACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTG 2051
             |||||||||| |||||||||||||||||||||||| |||||||||| ||||||||||||
Db        301 TACATGAAGCTGCAGCAAAAGGAAAATATGAAATATGCAAACTTCTACTCCAGCATGGTG 360

Qy       2052 CAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAG 2111
             |||||||||| || |||||||| ||||| |||||||| |||||||||||||||||||||
Db        361 CAGACCCTACAAAGAAAAACAGAGATGGTAATACTCCTCTGGATCTTGTTAAAGATGGAG 420

Qy       2112 ATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGG 2171
             |||||||||||||||||| ||||| || |||||||||||||| ||||||||||||||||
Db        421 ATACAGATATTCAAGATCTACTTAGAGGTGATGCAGCTTTGCTGGATGCTGCCAAGAAGG 480

Qy       2172 GTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCC 2231
             |||||||||||||||||||| |||||||||||||||||||||||||||||| |||||||
Db        481 GTTGTTTAGCCAGAGTGAAGAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATACCC 540

Qy       2232 AAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAG 2291
             |||| | ||||| |||||||||||||| |||||||||||||||| || | |||||| |
Db        541 AAGGTCGGCATTCTACACCTTTACATTTGGCAGCTGGTTATAATAACTTGGTAGTTGCTG 600

Qy       2292 AGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGA 2331
             ||||||| || ||||| |||||||||||||||||| ||
Db        601 AGTATTTATTGCAACATGGAGCTGATGTGAATGCCCAGGA 640
```

RESULT 35
BQ267594
LOCUS BQ267594 553 bp mRNA linear EST 15-JUL-2003
DEFINITION ij90c10.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:5778810 5' similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
ACCESSION BQ267594
VERSION BQ267594.1 GI:20492659
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 553)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute

```
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 429.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5778810"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact Hiroshi Inoue, MD/PhD for further
                     information on this library (Metabolism Division, Permutt
                     Laboratory, Washington University School of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

  Query Match            14.5%;  Score 549;  DB 13;  Length 553;
  Best Local Similarity  100.0%;  Pred. No. 4.7e-126;
  Matches  549;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 64

Qy       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 124

Qy       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 184

Qy       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 244

Qy       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 304

Qy       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 364

Qy       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 424

Qy       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 484

Qy       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 544

Qy       3781 TTGCTGAAA 3789
             |||||||||
Db        545 TTGCTGAAA 553


RESULT 36
CF729155
LOCUS       CF729155                691 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  UI-M-HD0-clc-f-23-0-UI.r1 NIH_BMAP_HD0 Mus musculus cDNA clone
            IMAGE:30615214 5', mRNA sequence.
ACCESSION   CF729155
VERSION     CF729155.1  GI:37603323
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousef1.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30615214"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HD0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH)."
ORIGIN

  Query Match            14.4%;  Score 545;  DB 14;  Length 691;
  Best Local Similarity  86.8%;  Pred. No. 4.9e-125;
  Matches  599;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy       2501 GCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTT 2560
             ||||||||||||| || ||||||||||||||| ||||||||||||||||||||||||||||
Db          2 GCCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTT 61

Qy       2561 TCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCT 2620
             ||||| |||||||||| || ||  |||||||||||||||||| |||||||||| |
Db         62 TCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACG 121

Qy       2621 TGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTC 2680
             || ||||| |||||||||||| ||| |||||  |||| ||||||||||||||||||||
```

```
Db        122 TGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTG 181

Qy       2681 TCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCT 2740

Db        182 TCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCT 241

Qy       2741 GGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGT 2800

Db        242 GGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGT 301

Qy       2801 TTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTT 2860

Db        302 TTGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTT 361

Qy       2861 GGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTT 2920

Db        362 GGACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTT 421

Qy       2921 GAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAA 2980

Db        422 GAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACAAG 481

Qy       2981 CTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACT 3040

Db        482 CTGATTAAAGGAGTTGAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAACT 541

Qy       3041 TTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTT 3100

Db        542 CTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGANTAAGAATTT 601

Qy       3101 CAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCA 3160

Db        602 CAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCA 661

Qy       3161 GGTGGAATCTTCAACAGATACAATATTCTC 3190

Db        662 GGCGGCGTCTTCAACAGATACAACATTCTC 691


RESULT 37
BG245969
LOCUS       BG245969                1099 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602358786F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4487204 5',
            mRNA sequence.
ACCESSION   BG245969
VERSION     BG245969.1  GI:12755784
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1099)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10331  row: 1  column: 21
            High quality sequence stop: 698.
FEATURES             Location/Qualifiers
     source          1. .1099
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4487204"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
ORIGIN

  Query Match           14.3%;  Score 544;  DB 12;  Length 1099;
  Best Local Similarity   88.5%;  Pred. No. 9.5e-125;
  Matches  680;  Conservative    0;  Mismatches   80;  Indels    8;  Gaps    8;

Qy       2918 GTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCAC 2977

Db          1 GTTGAAATGGGCCACAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACAC 60

Qy       2978 AAACTAATTAAAGGAGTCGAGAGACTTATCTC-CGGACAACAAGGTCTTAACCCATATTT 3036

Db         61 AAGCTGATTAAAGGAGTTGAAAGGCTGATCTCGTGGACAACAAGGTCTTAATCCATATTT 120

Qy       3037 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3096

Db        121 AACTCTAAACAACTCTGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGA 180

Qy       3097 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3156

Db        181 ATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCA 240

Qy       3157 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 3216

Db        241 CGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 300

Qy       3217 ACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGC 3276

Db        301 ATTGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGC 360

Qy       3277 CAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTT 3336

Db        361 AAACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATAAGGGCTT 420

Qy       3337 TGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAA 3396

Db        421 TGATGAAAGGCATGCATACATAGGTGGCATGTATGGAGCTGGAATTTATTTTGCTGAGAA 480

Qy       3397 CTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGT-GTCCAGTTC 3455

Db        481 CTCTTCCAAAAGCAATCAATATGTGTATGGAATCGGAGGTGGCACCGGATGGTCCAATTC 540

Qy       3456 ACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGG 3515

Db        541 ACAAAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCT-GTTTGTCGAGTAACCTTGG 599

Qy       3516 GAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACT 3575

Db        600 GAAAGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGGCATCACT 659

Qy       3576 CAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAG 3635

Db        660 CGGTCACTGGGCGG-CCAGTGTAAATGG-CTAGCATTAG-TGAATATG-TATTTACAGAG 715

Qy       3636 GAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGC 3683

Db        716 GAAAACAGG-TTATACTGAAAATCTAATTACTTTACCGATGGTAGGCC 762


RESULT 38
BE504478/c
LOCUS       BE504478                562 bp    mRNA    linear   EST 04-AUG-2000
DEFINITION  hz57a02.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3212042 3'
            similar to TR:O95271 O95271 TRF1-INTERACTING ANKYRIN-RELATED
            ADP-RIBOSE POLYMERASE. ;, mRNA sequence.
```

```
ACCESSION   BE504478
VERSION     BE504478.1  GI:9706886
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 503.
FEATURES             Location/Qualifiers
     source          1. .562
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3212042"
                     /tissue_type="carcinoid"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Lu24"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Plasmid DNA from the normalized
                     library NCI_CGAP_Lu5 was prepared, and ss circles were
                     made in vitro. Following HAP purification, this DNA was
                     used as tracer in a subtractive hybridization reaction.
                     The driver was PCR-amplified cDNAs from a pool of 5,000
                     clones made from the same library (cloneIDs
                     1414920-1417991 and 1520904-1522439). Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
ORIGIN

  Query Match           14.2%;  Score 538.6;  DB 10;  Length 562;
  Best Local Similarity   97.3%;  Pred. No. 1.9e-123;
  Matches  547;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       3224 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3283
             ||||||||||||||||||   |||  |  ||    |  ||| ||||||||||||||||||
Db        562 GAAAGATACACTCACCGGGAGAAAAGAAGTTTCTGAAGAAAACACAACCATGCCAATGAA 503

Qy       3284 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3343
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         502 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 443

Qy       3344 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3403
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         442 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 383

Qy       3404 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3463
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         382 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 323

Qy       3464 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3523
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         322 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 263

Qy       3524 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3583
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         262 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTNCAGGTCATCACTCAGTCACT 203

Qy       3584 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         202 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 143

Qy       3644 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         142 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 83

Qy       3704 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3763
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 23

Qy       3764 CTCTACGTTTTACTCCTTTGCT 3785
             |||||||||||||||||||||||
Db          22 CTCTACGTTTTACTCCTTTGCT 1


RESULT 39
CF538447
LOCUS       CF538447                621 bp    mRNA    linear   EST 12-SEP-2003
DEFINITION  UI-M-GI0-che-m-16-0-UI.r1 NIH_BMAP_GI0 Mus musculus cDNA clone
            IMAGE:30533583 5', mRNA sequence.
ACCESSION   CF538447
VERSION     CF538447.1  GI:34590429
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30533583"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_GI0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is AGCGAGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN

  Query Match           14.2%;  Score 537.8;  DB 14;  Length 621;
  Best Local Similarity   91.6%;  Pred. No. 3.1e-123;
```

```
  Matches  569;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGACCC 60

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              |||||||||||| || || |||||||| ||||||||||||||||  | ||||||||||| 
Db         61 ATCTGCCAAAGCTGTTCTGACTGGTGACTATAAGAAAGATGAACTTCTGGAAAGTGCCAG 120

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              |||||||||||| |||||||||||||| || || |||||||| |||||||||||||| ||
Db        121 GAGTGGCAATGAGGAAAAAATGATGGCCCTGCTTACACCATTGAATGTCAACTGCCATGC 180

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              |||||||||||||||||||||||||||| |||||||||||||||| |||||||||| ||
Db        181 AAGTGATGGCAGAAAGTCAACTCCATTGCATTTGGCAGCAGGATACAACAGAGTAAAAAT 240

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              |||||| |||||||||| ||||| |||||||||||||||||||||||| ||||||||
Db        241 TGTACAACTGTTACTGCACCATGGGGCTGATGTCCATGCTAAAGATAAAGGGGATCTGGT 300

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              |||| |||||||||||||| ||||| |||||||||||||||||||||||| ||||||||||
Db        301 ACCACTACACAATGCCTGCTCTTACGGTCATTATGAAGTAACTGAACTTCTGGTCAAGCA 360

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              |||||||||||| |||||||||||| ||||||| |||||||||||||||||||||||||
Db        361 CGGTGCCTGTGTGAATGCAATGGACCTGTGGCAGTTCACTCCTCTTCATGAGGCAGCTTC 420

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              |||||||||| |||||||||||||||||| ||| ||||||| || ||||| |||||||||
Db        421 TAAGAACAGGATTGAAGTATGTTCTCTCCTCCTAAGTTACGGGGCAGATCCAACCCTGCT 480

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              || |||||||||||||||||||||||||| |||| |||||  ||||| | || ||||||||
Db        481 GAACTGTCACAATAAAAGTGCTATAGACCTGGCCCCCACGGCACAGCTGAAGGAAAGATT 540

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              | | |||||||||||||||| || ||||||||||||||||||||||||||||||  |
Db        541 ATCGTATGAATTTAAAGGCCATTCATTGCTGCAAGCTGCACGAGAAGCTGATGTTACAAG 600

Qy       1243 AATCAAAAAACATCTCTCTCT 1263
              ||||||||||||||| |||||
Db        601 AATCAAAAAACATCTTTCTCT 621


RESULT 40
CF734609
LOCUS       CF734609                 703 bp    mRNA    linear   EST 10-OCT-2003
DEFINITION  UI-M-HB0-ckg-a-02-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
            IMAGE:30546721 5', mRNA sequence.
ACCESSION   CF734609
VERSION     CF734609.1  GI:37630945
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30546721"
                     /tissue_type="whole eye"
                     /dev_stage="embryo 12.5,13.5,14.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="NIH_BMAP_HB0"
                     /note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel.First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site .Double strand cDNA was
                     size selected according to mRNA size fraction ,ligated
                     with EcoR I adaptor , digested with NotI and then cloned
                     directionally into pYX-Asc vector . The library tag
                     sequence located between the Not I site and the polyA tail
                     is TTATTGAAGT. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH)."
ORIGIN

  Query Match           14.2%;  Score 537.6;  DB 14;  Length 703;
  Best Local Similarity   86.7%;  Pred. No. 3.5e-123;
  Matches  602;  Conservative    0;  Mismatches   91;  Indels    1;  Gaps    1;

Qy       2572 TGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCC 2631
              |||||| || ||  |||||||||||||||||||| ||||||||||  | || ||||| ||
Db          1 TGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGCTACAAACC 60

Qy       2632 TCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCC 2691
              |||||||||| ||| |||||  |||| |||||||||||||||||||| ||||||||||
Db         61 TCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTGTCTTCAGGTCC 120

Qy       2692 ATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTC 2751
              || |||||||| ||||| |||||||||||||| || |||||||||||||| || || ||
Db        121 GTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTGGCAGCTTCTC 180

Qy       2752 AGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAA 2811
              |||||||| |||| |||||||||||||| |||| |||||| |  ||||| | | |||
Db        181 GGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTTTGCAAAGAAA 240

Qy       2812 GGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCA 2871
              |||| | ||||| | |||||||| |||||||| ||| ||||||| ||||||||||||||
Db        241 AGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCA 300

Qy       2872 CCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCA 2931
              | | |||||||||||||||||||||||||| || || |||||  ||||||| ||||| ||
Db        301 CTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCA 360

Qy       2932 CAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGG 2991
              |||||| ||||| |||||||||||||||||| ||||||| | ||||| || ||||||||
Db        361 CAAGGAACTGAAAGAGATTGGAATCAATGCCTATGGACATCGACACAAGCTGATTAAAGG 420

Qy       2992 AGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTC 3051
              ||| || || || ||||| |||||||||||||||| |||||||||||  | |||| |||
Db        421 AGTTGAAAGGCTGATCTCTGGACAACAAGGTCTTAATCCATATTTAACTCTAAACAACTC 480

Qy       3052 TGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGA 3111
              |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||
Db        481 TGGTAGTGGAACAATTCTCATAGATCTGTCTCCTGATGATAAAGAATTTCAGTCTGTGGA 540

Qy       3112 GGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTT 3171
              |||||||||| ||||| || || |||||||||||||| ||||| ||||| ||  ||||
Db        541 AGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATGGTGGTCACGCAGGCGGCGTCTT 600

Qy       3172 CAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATA 3231
```

```
                 || |||||||| |||||||||||||||||||||||||| |||| | ||||||||||
Db        601 CANCAGATACAACATTCTCAAGATTCAGAAGGTTTGTAACANGAAATTGTGGGAAAGATA 660

Qy       3232 CACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3265
                || |||||||||||| ||||||||||||||||
Db        661 TACACACCGGAGAAAAG-AGTTTCTGAAGAAAAC 693


RESULT 41
BU570530
LOCUS       BU570530                765 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10405516 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6622292
            5', mRNA sequence.
ACCESSION   BU570530
VERSION     BU570530.1  GI:22920830
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2875  row: f  column: 20
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6622292"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             14.1%;  Score 535.8;  DB 13;  Length 765;
  Best Local Similarity   87.4%;  Pred. No. 1e-122;
  Matches  634;  Conservative    0;  Mismatches   29;  Indels   62;  Gaps    2;

Qy       1076 GAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAAT 1135
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAAT 62

Qy       1136 AAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTT 1195
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTT 122

Qy       1196 AAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACAT 1255
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACAT 182

Qy       1256 CTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATT------- 1308
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        183 CTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGATAGAG 242

Qy       1309 ----------------------------------------------------------GCATTG 1314
                                                                          ||||||
Db        243 ATGCTGCAGCCTGAAACTCCAGATTATATCATTTTCCACCACTGATTTTGCAGAGCATTG 302

Qy       1315 TGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGG 1374
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGG 362

Qy       1375 AGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAA 1434
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAA 422

Qy       1435 AGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGA 1494
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 AGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGA 482

Qy       1495 TAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCG 1554
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCG 542

Qy       1555 CCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTT 1614
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 NCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTT 602

Qy       1615 ACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGT-AATT 1673
                |||||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||
Db        603 ACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAAGGTATCTCATTAGGTAAATT 662

Qy       1674 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1733
                |||||||||||||||||    |  ||    ||  ||    ||  ||  |||
Db        663 CAGAGGCAGACAGACAATTTGCTGGGAAAGCTGCCAAAGGGCCTGGNAGAATGGTCCAAA 722

Qy       1734 AACTG 1738
                |||||
Db        723 AACTG 727


RESULT 42
CF726657
LOCUS       CF726657                679 bp    mRNA    linear   EST 09-OCT-2003
DEFINITION  UI-M-HB0-ckh-n-10-0-UI.r1 NIH_BMAP_HB0 Mus musculus cDNA clone
            IMAGE:30547809 5', mRNA sequence.
ACCESSION   CF726657
VERSION     CF726657.1  GI:37600825
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30547809"
```

/tissue_type="whole eye"
/dev_stage="embryo 12.5,13.5,14.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HB0"
/note="Organ: Eye; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel.First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site .Double strand cDNA was size selected according to mRNA size fraction ,ligated with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTATTGAAGT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Query Match 14.1%; Score 535.6; DB 14; Length 679;
Best Local Similarity 86.7%; Pred. No. 1.1e-122;
Matches 589; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```
Qy       2335 AGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCT 2394
             ||||||||| ||||||||||||||| |||||||| || || ||||||||||||||||||| |
Db             1 AGGAGGACTCATTCCTTTACATAACGCAGCATCCTATGGACATGTAGATGTAGCAGCTTT 60

Qy       2395 ACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCA 2454
              || |||||||| ||||| ||||||||||||||||||||||||||||||||||| || ||
Db            61 GCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTTTCACACCGTTACA 120

Qy       2455 CGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGA 2514
              |||||||||||||||||| |||||||||||||||| ||| | ||||||||||||||
Db           121 TGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGGCCCATGGAGCTGA 180

Qy       2515 CCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGT 2574
              || |||||||||||||| |||||||||||||||||||||||||||||||| |||||||
Db           181 TCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGT 240

Qy       2575 CAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCA 2634
              ||| || ||  |||||||||||||||||| |||||||||||  | || ||||| |||||
Db           241 CAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGTGCTACAAACCTCA 300

Qy       2635 AGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATC 2694
              |||||||| ||| |||||   |||| |||||||||||||||||| ||||||||||| ||
Db           301 AGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTGTCTTCAGGTCCGTC 360

Qy       2695 TAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGA 2754
              |||||||| ||||| |||||||||||||| || ||||||||||||||| || || || ||
Db           361 CAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTGGCAGCTTCTCGGA 420

Qy       2755 ACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGA 2814
              ||||||  |||| |||||||||||||| | |||| |||||| |  ||||| | | ||| ||
Db           421 ACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTTTGCAAAGAAAAGA 480

Qy       2815 GGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCT 2874
              || | ||||| | |||||||| |||||||| ||| ||||||| ||||||||||||||| |
Db           481 GGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTGGACTTGAGCACTT 540

Qy       2875 AATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAA 2934
              |||||||||||||||||||||||||| || || |||||  ||||||| ||||| ||||
Db           541 GATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTGAAATGGGCCACAA 600

Qy       2935 GGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGT 2994
              ||| ||||| |||||||| |||||||| ||||||||| | ||||| || ||||||||||
Db           601 GGAACTGAAAGAGATTGGNATCAATGCCTATGGACATCGACACAAGCTGATTAAAGGAGT 660

Qy       2995 CGAGAGACTTATCTCCGGA 3013
              || || || ||||| |||
Db           661 TGAAAGGCTGATCTCTGGA 679
```

RESULT 43
BF585334
LOCUS BF585334 996 bp mRNA linear EST 12-DEC-2000
DEFINITION 602102008F2 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4225393 5', mRNA sequence.
ACCESSION BF585334
VERSION BF585334.1 GI:11658962
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 996)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM9816 row: p column: 02
High quality sequence stop: 686.
FEATURES Location/Qualifiers
source 1. .996
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4225393"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Co24"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 14.1%; Score 534.4; DB 10; Length 996;
Best Local Similarity 83.6%; Pred. No. 2.4e-122;
Matches 653; Conservative 0; Mismatches 122; Indels 6; Gaps 4;

```
Qy       2325 CCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGT---AG 2381
              |||| |||||||||||||| |||||||||||||| ||||||| || || |||||   |
Db             1 CCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTATGGACATGTATGAT 60

Qy       2382 ATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTT 2441
                |  | || |||||||| ||||| ||||||||||||||||||||||||||||||
Db            61 GTAAGCAGACTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGACAAATGGGCTT 120

Qy       2442 TCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAG 2501
              ||||||| || || |||||||||||||||||| |||||||||||||||| ||| | |
Db           121 TCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGCTTTATTGTTGG 180

Qy       2502 CCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTT 2561
              |||||||||||||| || |||||||||||||| |||||||||||||||||||||||||||
Db           181 CCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTTAGATTTAGTTT 240

Qy       2562 CAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTT 2621
              |||| |||||||||||  || ||  |||||||||||||||||| ||||||||||  | |
Db           241 CAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGCTCTGCCTACGT 300

Qy       2622 GTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCT 2681
              | ||||| |||||||||||| ||| |||||  |||| ||||||||||||||||||| |
Db           301 GCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGCAGATGCTCTGT 360

Qy       2682 CTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTG 2741
              ||||||||| || |||||||| ||||| |||||||||||||| || ||||||||||||
Db           361 CTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGACAACTTATCTG 420
```

```
Qy       2742 GGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTT 2801
Db        421 GCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGGTGCTACTGGTT 480

Qy       2802 TGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTG 2861
Db        481 TGCAAAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCAGTTCATAAGGAACCTTG 540

Qy       2862 GACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTG 2921
Db        541 GACTTGAGCACTTGATGGATATATTTGAGAGAGAACAGATTACCTTAGATGTCCTAGTTG 600

Qy       2922 AGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAAC 2981
Db        601 AAATGGGCCACAAGGAACTGAAAGAGATTGGAATC-ATGCCTATGGACATCGACACCAAG 659

Qy       2982 TAATTAAAGGAGTCGAGAGACTTATCTCCGGACAAC-AAGGTCTTAACCCATATTTAACT 3040
Db        660 CTGATTAAGGAGTTGAAAGGCTGATCTCTGGACAACAAAGGTCTTAATCCATATTTTACT 719

Qy       3041 TTGAACACCTCTGGTAGTGGAACAATTCTTAT-AGATCTGTCTCCTGATGATAAAGAGTT 3099
Db        720 CTTAACAAATCTGGTAGTGGGACACATCTCATAAAATCTGTGTCCCGCGGATAANGAATC 779

Qy       3100 T 3100
Db        780 T 780
```

RESULT 44
CA317519
LOCUS CA317519 769 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-FW0-cbn-m-24-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone IMAGE:6812065 5', mRNA sequence.
ACCESSION CA317519
VERSION CA317519.1 GI:24535643
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 769)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .769
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6812065"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FW0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN

Query Match 14.0%; Score 533.2; DB 14; Length 769;
Best Local Similarity 91.4%; Pred. No. 4.5e-122;
Matches 565; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```
Qy       3090 ATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATG 3149
Db          1 ATAAAGAATTTCAGTCTGTGGAAGAAGAGATGCAGAGTACTGTGCGGGAGCACAGAGATG 60

Qy       3150 GAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTA 3209
Db         61 GTGGTCACGCAGGCGGCGTCTTCAACAGATACAACATTCTCAAGATTCAGAAGGTTTGTA 120

Qy       3210 ACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACA 3269
Db        121 ACAAGAAATTGTGGGAAAGATATACACACCGGAGAAAAGAAGTTTCTGAAGAAAACCACA 180

Qy       3270 ACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACA 3329
Db        181 ACCATGCAAACGAAAGGATGTTATTTCATGGGTCTCCTTTTGTGAATGCGATTATCCATA 240

Qy       3330 AAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTG 3389
Db        241 AGGGCTTTGATGAAAGGCATGCATACATAGGTGGCATGTTTGGAGCTGGAATTTATTTTG 300

Qy       3390 CTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTC 3449
Db        301 CTGAGAACTCTTCCAAAAGCAATCAATATGTGTATGGAATTGGAGGTGGCACCGGATGTC 360

Qy       3450 CAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAA 3509
Db        361 CAATTCACAAAGACAGATCGTGTTACATTTGTCACAGGCAGCTGCTGTTTTGTCGAGTAA 420

Qy       3510 CCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTC 3569
Db        421 CCTTGGGAAAGTCTTTCTTGCAGTTCAGCGCAATGAAAATGGCACATTCTCCTCCCGGCC 480

Qy       3570 ATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTT 3629
Db        481 ATCACTCGGTCACTGGCCGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTT 540

Qy       3630 ACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3689
Db        541 ACAGAGGAGAACAGGCTTATCCTGAATATCTAATTACTTACCAGATTGTAAGGCCTGAAG 600

Qy       3690 GTATGGTCGATGGATAAA 3707
Db        601 GTATGGTTGATGGCTGAA 618
```

RESULT 45
CD350434
LOCUS CD350434 655 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-M-FY0-cfs-1-03-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone IMAGE:6855412 5', mRNA sequence.
ACCESSION CD350434
VERSION CD350434.1 GI:31141949
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 655)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES Location/Qualifiers
source 1. .655
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6855412"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN

Query Match 13.9%; Score 528.2; DB 14; Length 655;
Best Local Similarity 87.9%; Pred. No. 7.8e-121;
Matches 575; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy       2191 GAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACC 2250
             ||| |||||||||||||||||||||||||||| |||||||||||  | ||||| |||||
Db          1 GAAATTGTCTTCTCCTGATAATGTAAATTGCCGTGATACCCAAGGTCGGCATTCTACACC 60

Qy       2251 TTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGG 2310
             ||||||||| |||||||||||||||| || |||||||| |||||||| || ||||| ||
Db         61 TTTACATTTGGCAGCTGGTTATAATAACTTGGAAGTTGCTGAGTATTTATTGCAACATGG 120

Qy       2311 AGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTA 2370
             |||||||||||||||| |||||||||||| |||||||||||| |||||||| ||
Db        121 AGCTGATGTGAATGCCCAGGACAAAGGAGGACTCATTCCTTTACATAACGCAGCATCGTA 180

Qy       2371 CGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGA 2430
               || |||||||||||||||| | || |||||||| ||||| ||||||||||||||||
Db        181 TGGACATGTAGATGTAGCAGCTTTGCTGATAAAGTACAATGCGTGTGTCAATGCCACGGA 240

Qy       2431 CAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGC 2490
             |||||||||||||||| || || ||||||||||||||||||| ||||||||||||||
Db        241 CAAATGGGCTTTCACACCGTTACATGAAGCAGCCCAAAAGGGACGGACACAGCTTTGTGC 300

Qy       2491 TTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTT 2550
             ||| ||| | |||||||||||||| || |||||||||||||| ||||||||||||||||
Db        301 TTTATTGTTGGCCCATGGAGCTGATCCTACTCTTAAAAATCAAGAAGGACAAACACCTTT 360

Qy       2551 AGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGC 2610
             |||||||||||||| |||||||||| || ||  |||||||||||||||||||| |||||
Db        361 AGATTTAGTTTCAGCAGATGATGTCAGTGCACTCTTGACAGCAGCCATGCCCCCCTCTGC 420

Qy       2611 TCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGC 2670
             ||||||  | || |||| |||||||||||||| ||| |||||  |||| ||||||||||
Db        421 TCTGCCTACGTGCTACAAACCTCAAGTGCTCAGTGGCGTGAGGGGCCCCGGAGCCACTGC 480

Qy       2671 AGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGA 2730
             ||||||||| |||||||||| || |||||||| |||||  |||||||||||||| || ||
Db        481 AGATGCTCTGTCTTCAGGTCCGTCCAGCCCATCCAGCCTCTCTGCAGCCAGCAGCCTCGA 540

Qy       2731 CAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGG 2790
             |||||||||||| || || || ||||||||  |||| ||||||||||||| | |||| ||
Db        541 CAACTTATCTGGCAGCTTCTCGGAACTGTCCGCAGTGGTTAGTTCAAGTGCAGCAGAAGG 600

Qy       2791 TGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCA 2844
             |||| |  ||||| | | ||| |||| | ||||| | ||||||||  ||||||||
Db        601 TGCTACTGGTTTGCANAGAAAAGAGGATTCAGGAATCGATTTTAGTATAACTCA 654
```

Search completed: March 30, 2004, 01:39:53
Job time : 9088.83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 01:40:08 ; Search time 1282.29 Seconds
(without alignments)
11024.221 Million cell updates/sec

Title: US-09-843-159B-1
Perfect score: 3797
Sequence: 1 ctttgaagacactggatttc.....cctttgctgaaaaaaaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3388.8	89.2	4275	10	US-09-972-115A-5	Sequence 5, Appli
2	3388.8	89.2	5778	14	US-10-163-587A-14	Sequence 14, Appl
3	3387.2	89.2	4127	14	US-10-199-937-100	Sequence 100, App
4	3387.2	89.2	4406	14	US-10-199-937-106	Sequence 106, App
5	3387.2	89.2	4992	14	US-10-199-937-134	Sequence 134, App
6	3387.2	89.2	5002	14	US-10-199-937-132	Sequence 132, App
7	3387.2	89.2	5075	12	US-10-302-172-556	Sequence 556, App
8	3387.2	89.2	6018	10	US-09-849-602-11	Sequence 11, Appl
9	3297	86.8	3400	9	US-09-509-196A-1	Sequence 1, Appli
10	3296.2	86.8	3508	14	US-10-199-937-1	Sequence 1, Appli
11	2969.6	78.2	4493	10	US-09-972-115A-1	Sequence 1, Appli
12	2968.2	78.2	4297	10	US-09-972-115A-3	Sequence 3, Appli
13	2439.8	64.3	2971	14	US-10-199-937-92	Sequence 92, Appl
14	2439.8	64.3	3353	14	US-10-199-937-88	Sequence 88, Appl
15	2357.8	62.1	3799	14	US-10-199-937-90	Sequence 90, Appl
16	1813.4	47.8	3984	14	US-10-199-937-3	Sequence 3, Appli
17	1813.4	47.8	4134	9	US-09-841-835-1	Sequence 1, Appli
18	1813.4	47.8	4134	10	US-09-972-115A-7	Sequence 7, Appli
19	1533.8	40.4	4435	14	US-10-252-157-342	Sequence 342, App
20	1449.6	38.2	4491	9	US-09-841-835-7	Sequence 7, Appli
21	1359.4	35.8	1568	14	US-10-199-937-98	Sequence 98, Appl
22	1297.8	34.2	4657	9	US-09-841-835-9	Sequence 9, Appli
23	1240.4	32.7	1466	12	US-10-276-774-340	Sequence 340, App
24	1168.4	30.8	3308	14	US-10-199-937-177	Sequence 177, App
25	1084.8	28.6	1103	14	US-10-199-937-159	Sequence 159, App
26	1008.8	26.6	5482	14	US-10-199-937-138	Sequence 138, App
27	994	26.2	2409	9	US-09-964-899-40	Sequence 40, Appl
28	794.4	20.9	797	14	US-10-252-157-341	Sequence 341, App
29	574	15.1	582	14	US-10-199-937-45	Sequence 45, Appl
c 30	518.4	13.7	520	10	US-09-972-115A-10	Sequence 10, Appl
31	467.6	12.3	1069	9	US-09-833-381-841	Sequence 841, App
32	461	12.1	465	9	US-09-833-381-1153	Sequence 1153, Ap
33	459	12.1	460	14	US-10-106-698-2223	Sequence 2223, Ap
34	459	12.1	460	14	US-10-199-937-5	Sequence 5, Appli
35	396.8	10.5	402	14	US-10-181-447A-11	Sequence 11, Appl
36	380.4	10.0	384	10	US-09-972-115A-9	Sequence 9, Appli
37	368	9.7	441	10	US-09-918-995-17041	Sequence 17041, A
38	359.6	9.5	1691	14	US-10-199-937-41	Sequence 41, Appl
39	359.6	9.5	1692	14	US-10-199-937-43	Sequence 43, Appl
40	356	9.4	356	14	US-10-199-937-49	Sequence 49, Appl
41	354.6	9.3	564	14	US-10-199-937-7	Sequence 7, Appli
42	346	9.1	392	14	US-10-199-937-14	Sequence 14, Appl
c 43	340.4	9.0	353	14	US-10-199-937-12	Sequence 12, Appl
44	323	8.5	353	13	US-10-040-739-1179	Sequence 1179, Ap
45	285.6	7.5	471	12	US-10-085-783A-34884	Sequence 34884, A

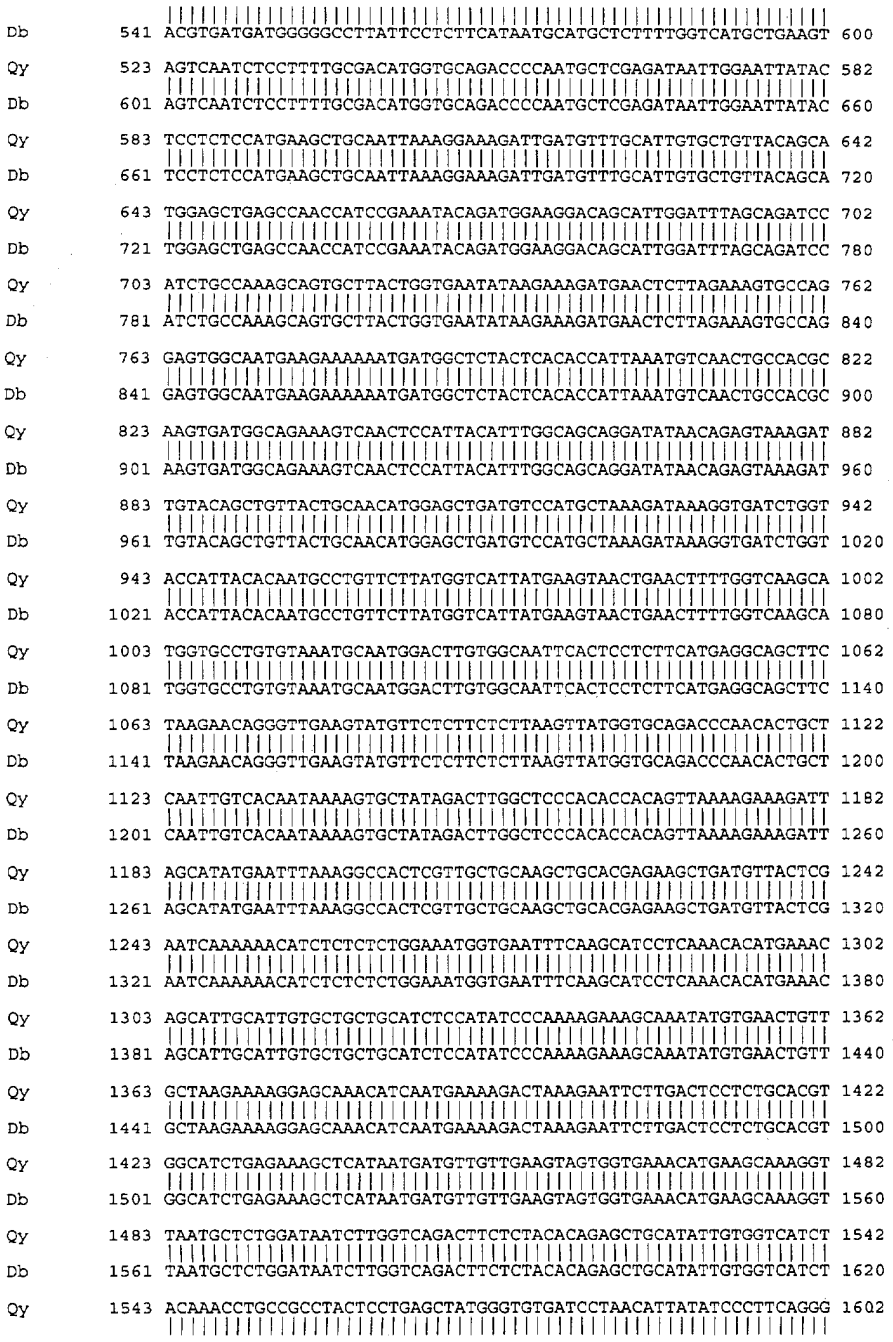
ALIGNMENTS

RESULT 1
US-09-972-115A-5
; Sequence 5, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 4275
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (284)..(3781)
; OTHER INFORMATION:
US-09-972-115A-5

Query Match 89.2%; Score 3388.8; DB 10; Length 4275;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
            ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db      481 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 540

Qy      463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
```


Db 541 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 600
Qy 523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db 601 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 660
Qy 583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db 661 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 720
Qy 643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db 721 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 780
Qy 703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db 781 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 840
Qy 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db 841 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 900
Qy 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 901 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 960
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db 961 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1020
Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 1021 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1080
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 1081 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1140
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1141 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1200
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1201 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1260
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1261 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1320
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1321 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1380
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1381 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1440
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1441 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1500
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1501 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1560
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1561 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1620
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1621 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1680
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1681 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1740
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1741 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1800
Qy 1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1801 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1860
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1861 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1920
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1921 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1980
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1981 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2040
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2041 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2100
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 2101 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2160
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2161 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2220
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2221 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2280
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2281 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2340
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2341 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2400
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2401 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2460
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2461 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2520
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2521 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2580
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2581 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2640
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2641 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2700
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2701 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2760

```
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2820

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2880

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2940

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3000

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3060

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3120

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3180

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3240

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3300

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3360

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3420

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3480

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3540

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3600

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3660

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3720

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3780

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3840

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3841 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 2
US-10-163-587A-14
; Sequence 14, Application US/10163587A
; Publication No. US20030096263A1
; GENERAL INFORMATION:
; APPLICANT: Oliveira, Marcos
; TITLE OF INVENTION: SELECTIVE PARP-1 TARGETING FOR DESIGNING CHEMO/RADIO SENSITIZING F
; FILE REFERENCE: 50229-306
; CURRENT APPLICATION NUMBER: US/10/163,587A
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 60/296,110
; PRIOR FILING DATE: 2001-06-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 5778
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3501)
; OTHER INFORMATION:
US-10-163-587A-14

Query Match 89.2%; Score 3388.8; DB 14; Length 5778;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

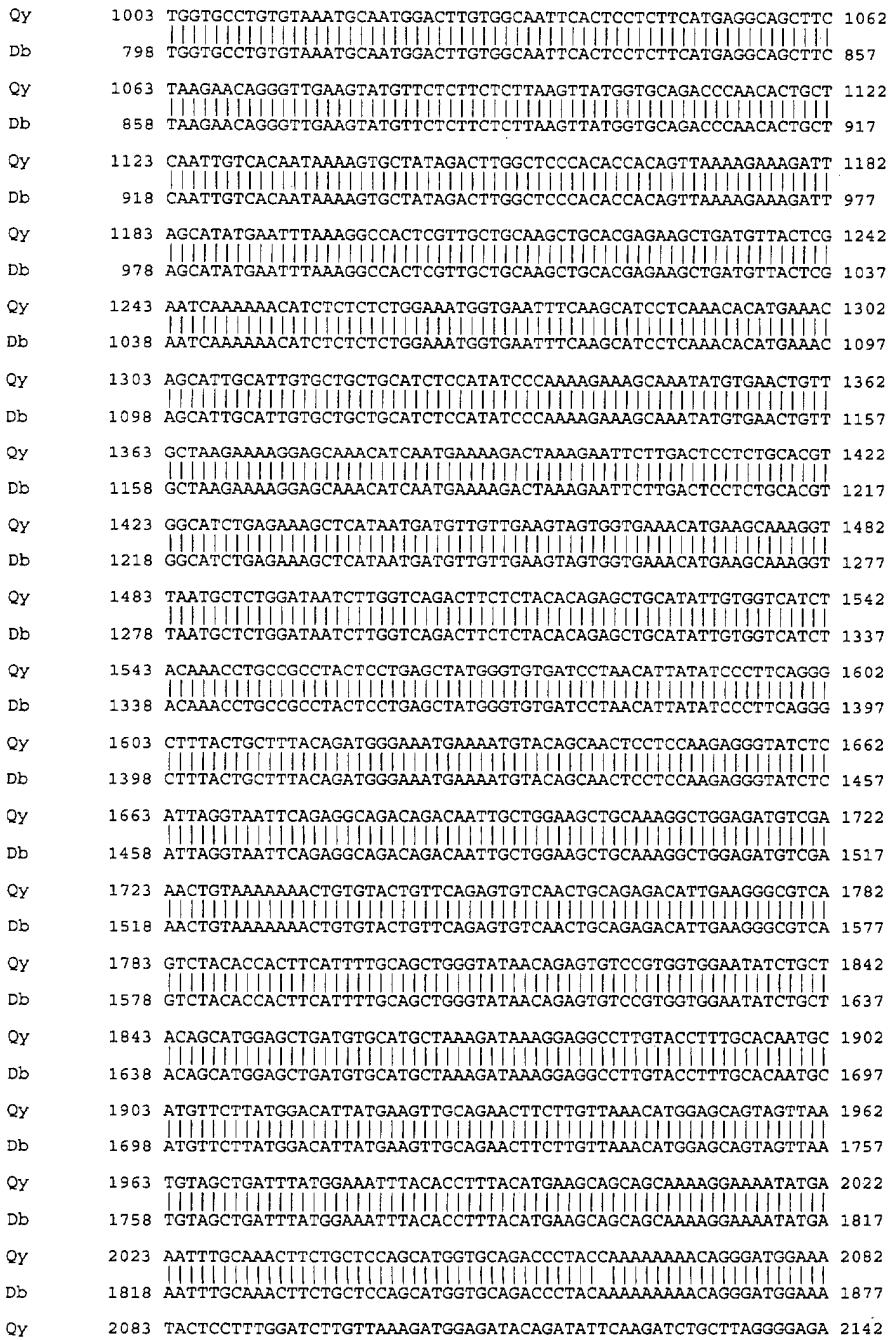
Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 557

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 617

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        618 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 677

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 737

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        738 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 797
```


Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 798 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 857
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 858 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 917
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 918 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 977
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1037
Qy 1243 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1038 AATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1097
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1098 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1157
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1158 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1217
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1218 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1277
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1278 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1337
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1338 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1397
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1398 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1457
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1458 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1517
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1518 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1577
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1578 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1637
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1638 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1697
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1698 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1757
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1758 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1817
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1877
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1878 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1937
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1938 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 1997
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 1998 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2057
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2058 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2117
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2118 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2177
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2178 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2237
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2238 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2297
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2358 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2418 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2477
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2478 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2537
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2538 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2597
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2657
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2658 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2717
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2718 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2777
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2778 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2837
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 2838 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2897
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 2898 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2957
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

```
Db       2958 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3017

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3018 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3077

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3078 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3137

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3138 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3197

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3198 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3257

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3258 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3317

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3318 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3377

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3378 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3437

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3438 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3497

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3498 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3557

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3558 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3589
```

RESULT 3
US-10-199-937-100
; Sequence 100, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 4127
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(3508)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (3509)..(4127)
US-10-199-937-100

Query Match 89.2%; Score 3387.2; DB 14; Length 4127;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

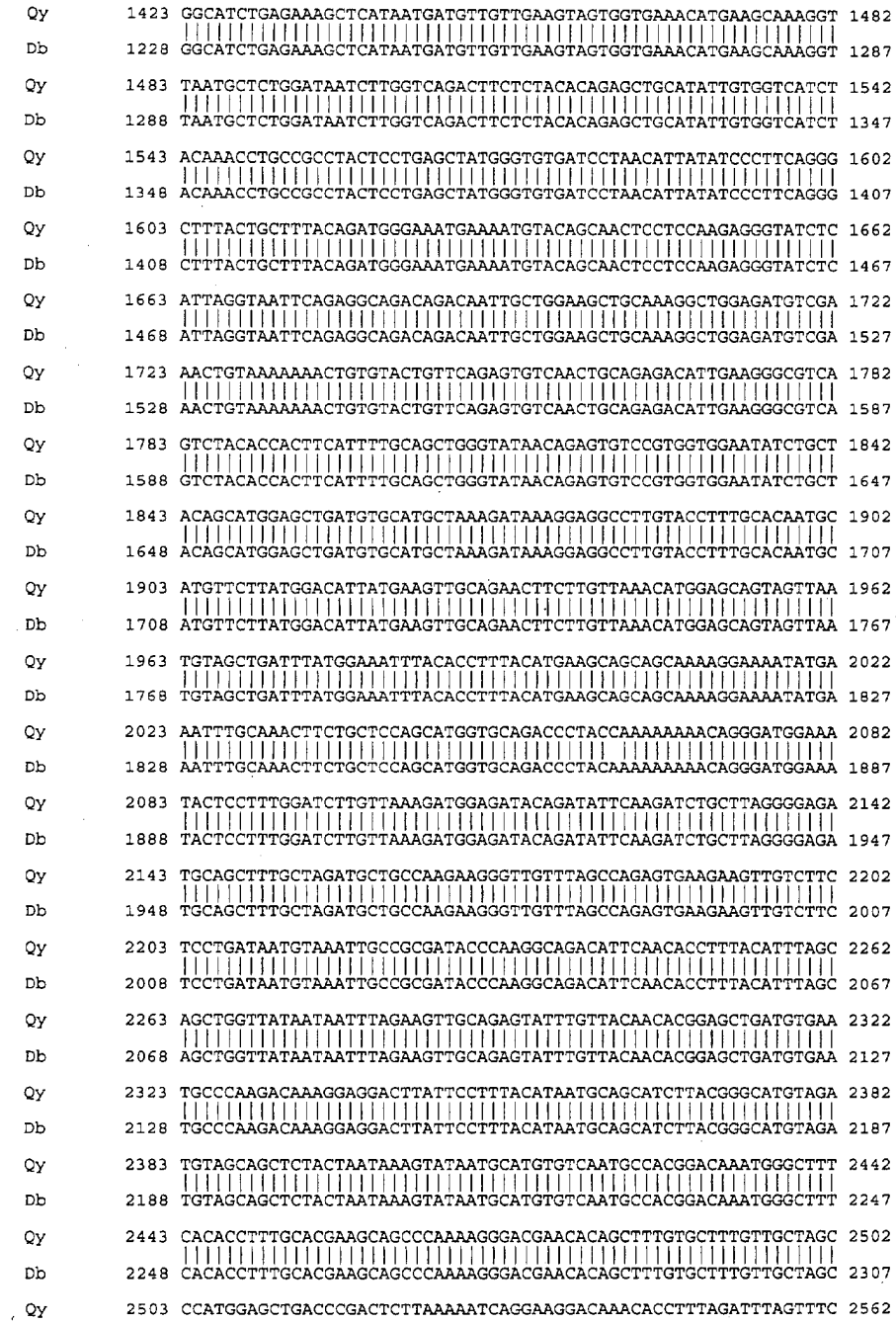
Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227
```


Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1587
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1588 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1647
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1648 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1707
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1708 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1767
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1768 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1827
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1887
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1888 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1947
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1948 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2007
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2008 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2067
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2068 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2127
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2188 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2247
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2248 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2307
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2308 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2427
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2428 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2487
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2488 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2547
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2607
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2608 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2728 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2787
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2847
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 2848 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2907
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 2908 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2967
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 2968 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3027
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3028 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3087
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3088 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3207
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3208 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3267
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3268 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3327
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3328 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3387
Qy 3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

```
Db       3388 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3447

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3507

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3508 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3567

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3568 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3599


RESULT 4
US-10-199-937-106
; Sequence 106, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
;   LENGTH: 4406
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(3787)
US-10-199-937-106

  Query Match             89.2%;  Score 3387.2;  DB 14;  Length 4406;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        487 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 546

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 606

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 666

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 726

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 786

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 846

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        847 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 906

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 966

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1026

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1086

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1146

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1206

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1266

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1267 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1326

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1327 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1386

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1387 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1446

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1447 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1506

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1507 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1566

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1567 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1626

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1627 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1686

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1687 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1746

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1747 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1806

Qy       1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1807 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1866

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1867 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1926

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1927 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1986
```

```
Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1987 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2046

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2047 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2106

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       2107 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2166

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2226

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2286

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2346

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2347 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2406

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2407 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2466

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2526

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2527 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2586

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2587 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2646

Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2647 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2706

Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2707 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2766

Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2767 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2826

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2827 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2886

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2887 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2946

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2947 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3006

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3007 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3066

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3067 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3126

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3127 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3186

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3187 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3246

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3247 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3306

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3307 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3366

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3367 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3426

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3427 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3486

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3487 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3546

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3547 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3606

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3607 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3666

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3667 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3726

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3727 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3786

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3787 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3846

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||||||||||||||
Db       3847 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3878
```

RESULT 5
US-10-199-937-134
; Sequence 134, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582

```
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
;   LENGTH: 4992
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (876)..(4373)
US-10-199-937-134

  Query Match            89.2%;  Score 3387.2;  DB 14;  Length 4992;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1073 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 1132

Qy       463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1133 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 1192

Qy       523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1193 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 1252

Qy       583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1253 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 1312

Qy       643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1313 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 1372

Qy       703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1373 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1432

Qy       763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1433 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1492

Qy       823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1493 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1552

Qy       883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1553 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1612

Qy       943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1613 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1672

Qy      1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1673 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1732

Qy      1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1733 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1792

Qy      1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1793 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1852

Qy      1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1853 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1912

Qy      1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1913 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1972

Qy      1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1973 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 2032

Qy      1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2033 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 2092

Qy      1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2093 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 2152

Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2153 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 2212

Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2213 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 2272

Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2273 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 2332

Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2333 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 2392

Qy      1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2393 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2452

Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2453 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2512

Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2513 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 2572

Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2573 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2632

Qy      1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2633 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2692

Qy      2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2693 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2752

Qy      2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2753 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2812

Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2813 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2872

Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2873 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2932

Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2933 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2992

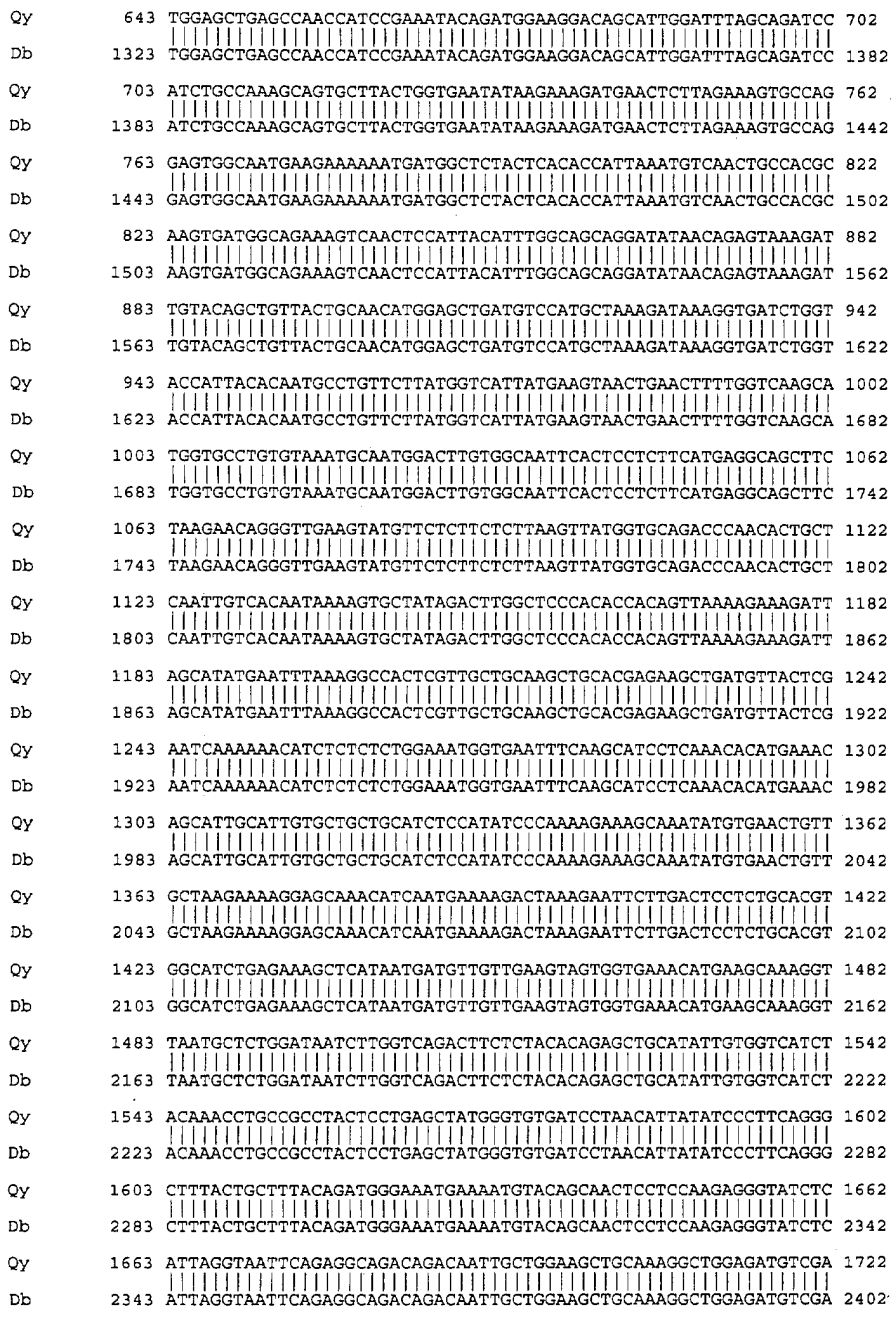
Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

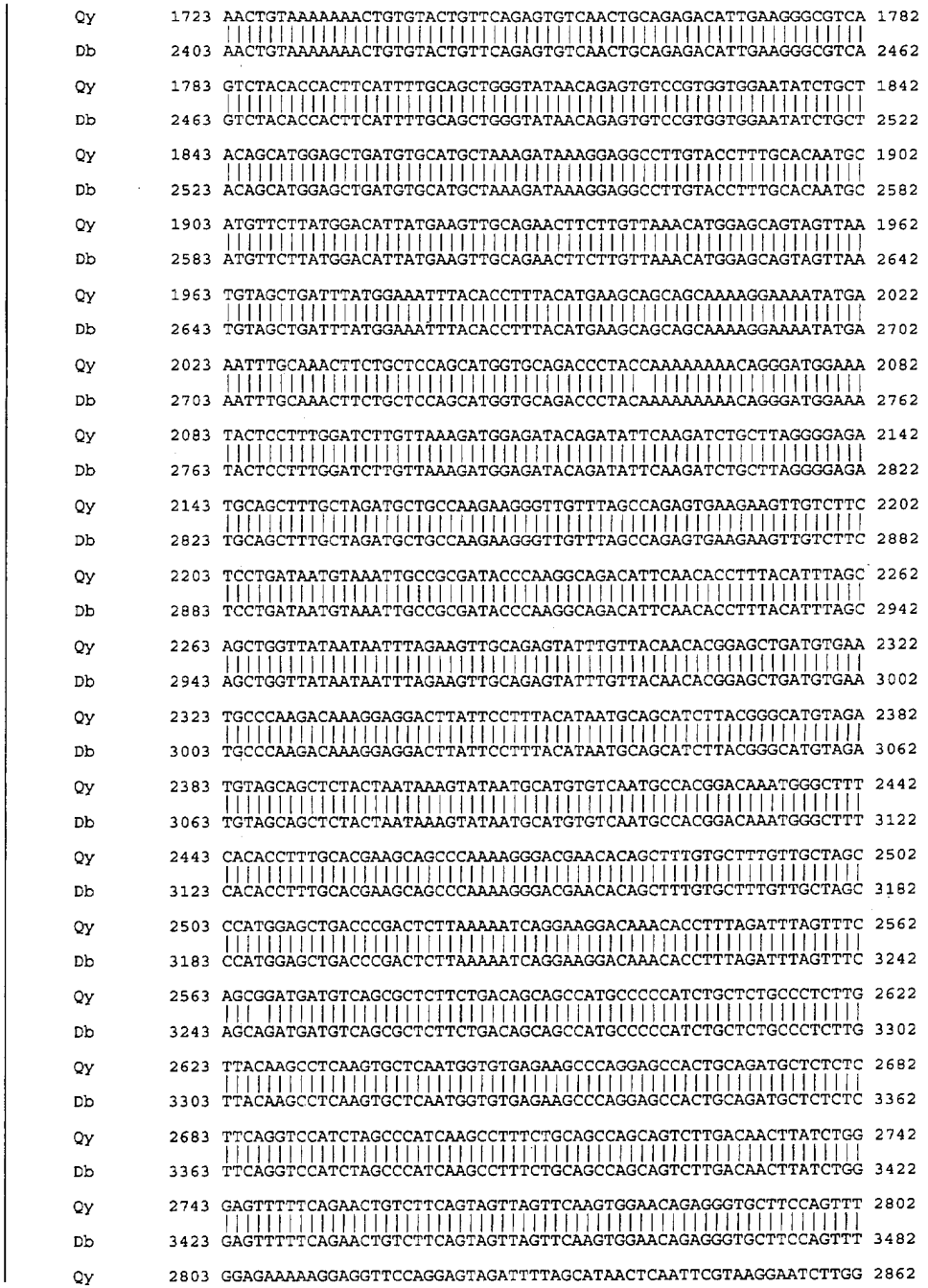
```
Db       2993 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 3052
Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       3053 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 3112
Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       3113 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 3172
Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       3173 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 3232
Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       3233 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 3292
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       3293 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 3352
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       3353 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 3412
Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       3413 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 3472
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       3473 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3532
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       3533 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3592
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       3593 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3652
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       3653 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3712
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       3713 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3772
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       3773 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3832
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       3833 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3892
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3893 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3952
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3953 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 4012
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       4013 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 4072
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       4073 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 4132
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       4133 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 4192
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       4193 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 4252
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       4253 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 4312
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       4313 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 4372
Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       4373 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 4432
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       4433 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4464
```

RESULT 6
US-10-199-937-132
; Sequence 132, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 132
; LENGTH: 5002
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)..(4383)
US-10-199-937-132

Query Match 89.2%; Score 3387.2; DB 14; Length 5002;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db       1083 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 1142
Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db       1143 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 1202
Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db       1203 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 1262
Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db       1263 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 1322
```


Qy 643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db 1323 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 1382
Qy 703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db 1383 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1442
Qy 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db 1443 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1502
Qy 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 1503 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1562
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db 1563 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1622
Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 1623 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1682
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 1683 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1742
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1743 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1802
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1803 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1862
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1863 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1922
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1923 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1982
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1983 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 2042
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 2043 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 2102
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 2103 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 2162
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 2163 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 2222
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 2223 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 2282
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 2283 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 2342
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 2343 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 2402


Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 2403 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2462
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 2463 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2522
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 2523 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 2582
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 2583 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2642
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2643 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2702
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 2703 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2762
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2763 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2822
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2823 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2882
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2883 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2942
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2943 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 3002
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 3003 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 3062
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 3063 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 3122
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 3123 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 3182
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 3183 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 3242
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 3243 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 3302
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 3303 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 3362
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 3363 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 3422
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 3423 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 3482
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862

```
Db       3483 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 3542

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922

Db       3543 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3602

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982

Db       3603 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3662

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042

Db       3663 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3722

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102

Db       3723 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3782

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162

Db       3783 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3842

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222

Db       3843 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3902

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

Db       3903 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3962

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342

Db       3963 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 4022

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402

Db       4023 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 4082

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462

Db       4083 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 4142

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

Db       4143 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 4202

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Db       4203 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 4262

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Db       4263 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 4322

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Db       4323 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 4382

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Db       4383 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 4442

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794

Db       4443 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4474


RESULT 7
US-10-302-172-556
; Sequence 556, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 803_1CNCP
;  CURRENT APPLICATION NUMBER: US/10/302,172
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/225,251
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: PCT US02/05095
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 09/799,451
;  PRIOR FILING DATE: 2001-03-05
;  NUMBER OF SEQ ID NOS: 950
;  SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 556
;   LENGTH: 5075
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (253)..(3750)
US-10-302-172-556

  Query Match             89.2%;  Score 3387.2;  DB 12;  Length 5075;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Db        450 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAAAATGGTGCAAATGTCCAAGC 509

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522

Db        510 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 569

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582

Db        570 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 629

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642

Db        630 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 689

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702

Db        690 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 749

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762

Db        750 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 809

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822

Db        810 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 869

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882

Db        870 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 929

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942

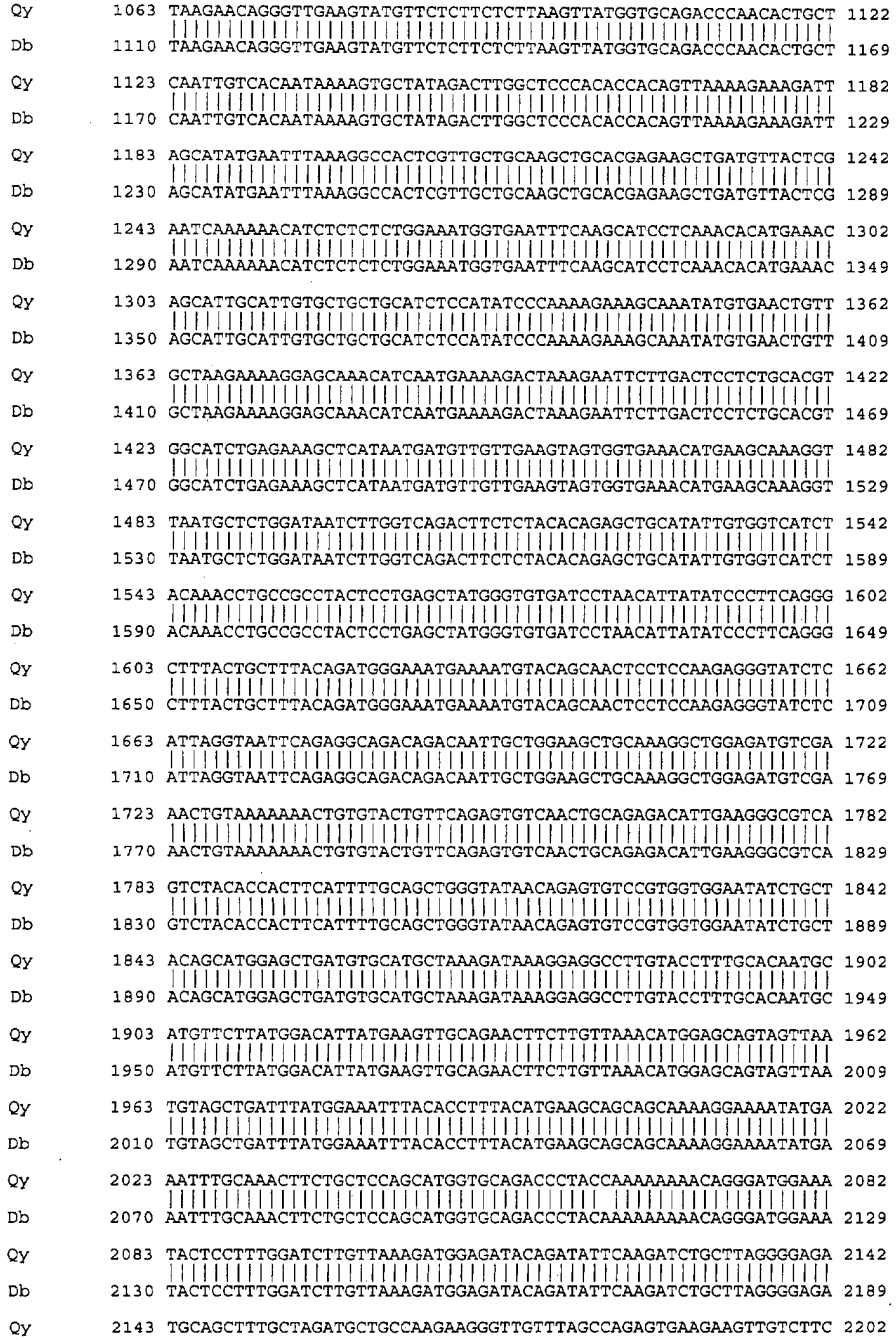
Db        930 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 989

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002

Db        990 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1049

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062

Db       1050 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1109
```


Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1110 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1169
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1170 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1229
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1230 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1289
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1290 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1349
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1350 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1409
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1410 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1469
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1470 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1529
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1530 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1589
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1590 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1649
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1650 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1709
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1710 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1769
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1770 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1829
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1830 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1889
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1890 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1949
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1950 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2009
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2010 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2069
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 2070 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2129
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2130 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2189
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2190 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2249
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2250 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2309
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2310 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2369
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2370 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2429
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2430 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2489
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2490 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2549
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2550 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2609
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2610 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2669
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2670 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2729
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2730 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2789
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2790 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2849
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2850 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2909
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2910 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2969
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2970 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3029
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3030 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3089
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3090 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3149
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3150 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3209
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3210 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3269
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282

```
Db       3270 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3329

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3389

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3449

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3450 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3509

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3569

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3629

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3689

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3749

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3750 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3809

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
              ||||||||||||||||||||||||||||||||
Db       3810 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3841


RESULT 8
US-09-849-602-11
; Sequence 11, Application US/09849602
; Publication No. US20030165834A1
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Matthew J.
;  APPLICANT:  Old, Lloyd J.
;  APPLICANT:  Stockert, Elisabeth
;  APPLICANT:  Chen, Yao-Tseng
;  TITLE OF INVENTION: Colon Cancer Antigen Panel
;  FILE REFERENCE: L0461/7105(JRV)
;  CURRENT APPLICATION NUMBER: US/09/849,602
;  CURRENT FILING DATE:  2001-05-04
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 6018
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-849-602-11

  Query Match             89.2%;  Score 3387.2;  DB 10;  Length 6018;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        418 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 477

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 537

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 597

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        598 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 657

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        658 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 717

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        718 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 777

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        778 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 837

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        838 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 897

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        898 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 957

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        958 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1017

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1018 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCAQTCCTCTTCATGAGGCAGCTTC 1077

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1078 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1137

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1197

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1257

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1317

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1377

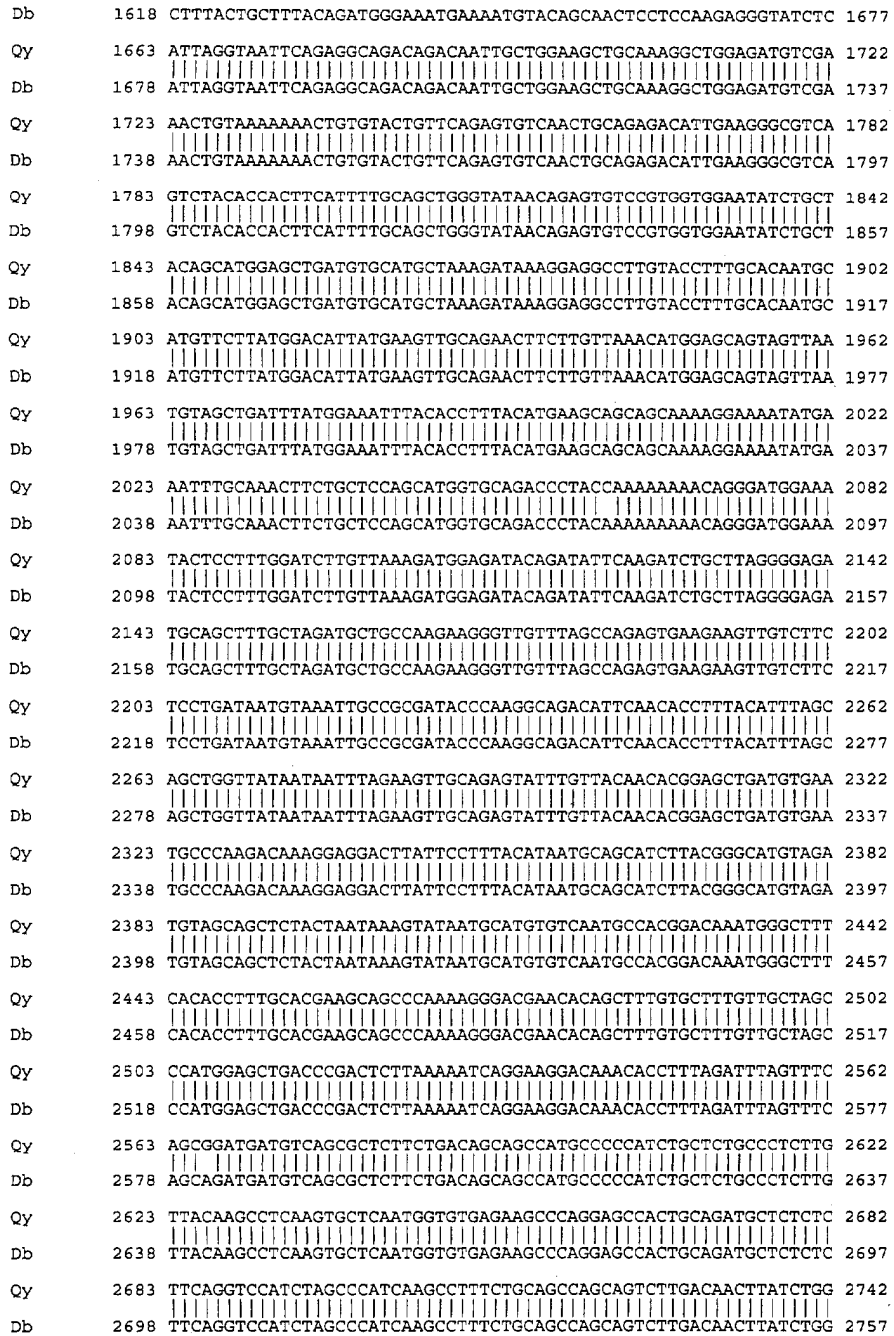
Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1437

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1497

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1557

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1617

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 1618 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1677
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1678 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1737
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1738 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1797
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1798 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1857
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1858 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1917
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1918 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1977
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1978 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2037
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2038 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2097
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2098 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2157
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2158 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2217
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2218 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2277
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2278 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2337
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2338 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2397
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2398 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2457
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2458 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2517
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2518 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2577
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2578 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2637
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2638 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2697
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2698 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2757
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2758 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2817
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2818 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2877
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2878 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2937
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2938 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2997
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 2998 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3057
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3058 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3117
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3118 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3177
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3178 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3237
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3238 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3297
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3298 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3357
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3358 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3417
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3418 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3477
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db 3478 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3537
Qy 3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db 3538 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3597
Qy 3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db 3598 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3657
Qy 3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db 3658 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3717
Qy 3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db 3718 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3777
Qy 3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db 3778 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3809

RESULT 9
US-09-509-196A-1
; Sequence 1, Application US/09509196A
; Patent No. US20020037582A1
; GENERAL INFORMATION:
; APPLICANT: DALY, Roger J.
; APPLICANT: SUTHERLAND, Robert L.
; TITLE OF INVENTION: A Potential Effector for the Grb7 Family of Signalling
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 1871-129
; CURRENT APPLICATION NUMBER: US/09/509,196A
; CURRENT FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: P09388
; PRIOR FILING DATE: 1997-09-23
; PRIOR APPLICATION NUMBER: PCT AU98/00795
; PRIOR FILING DATE: 1998-09-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-196A-1

Query Match 86.8%; Score 3297; DB 9; Length 3400;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy       482 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 541
Db         1 ATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGA 60

Qy       542 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 601
Db        61 CATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCA 120

Qy       602 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 661
Db       121 ATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATC 180

Qy       662 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 721
Db       181 CGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTT 240

Qy       722 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 781
Db       241 ACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAA 300

Qy       782 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 841
Db       301 ATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCA 360

Qy       842 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 901
Db       361 ACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAA 420

Qy       902 CATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 961
Db       421 CATGGACGTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGT 480

Qy       962 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCA 1021
Db       481 TCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGGCTGTGTAAATGCA 540

Qy      1022 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 1081
Db       541 ATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTA 600

Qy      1082 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGT 1141
Db       601 TGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTAAGAATAAAAGT 660

Qy      1142 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 1201
Db       661 GCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGC 720

Qy      1202 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 1261
Db       721 CACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCT 780

Qy      1262 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 1321
Db       781 CTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCT 840

Qy      1322 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 1381
Db       841 GCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAAC 900

Qy      1382 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 1441
Db       901 ATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCAT 960

Qy      1442 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1501
Db       961 AATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTT 1020

Qy      1502 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1561
Db      1021 GGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTC 1080

Qy      1562 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1621
Db      1081 CTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATG 1140

Qy      1622 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1681
Db      1141 GGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCA 1200

Qy      1682 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1741
Db      1201 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1260

Qy      1742 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1801
Db      1261 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1320

Qy      1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1861
Db      1321 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1380

Qy      1862 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1921
Db      1381 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTACGGACATTAT 1440

Qy      1922 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1981
Db      1441 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1500

Qy      1982 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2041
Db      1501 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 1560

Qy      2042 CAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 2101
Db      1561 CAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTT 1620

Qy      2102 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 2161
Db      1621 AAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCT 1680

Qy      2162 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 2221
Db      1681 GCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGC 1740

Qy      2222 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 2281
```

```
Db       1741 CGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTA 1800

Qy       2282 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 2341

Db       1801 GAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGA 1860

Qy       2342 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 2401

Db       1861 CTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATA 1920

Qy       2402 AAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 2461

Db       1921 AAGTATAATGCATCTCTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCA 1980

Qy       2462 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2521

Db       1981 GCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACT 2040

Qy       2522 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCT 2581

Db       2041 CTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCT 2100

Qy       2582 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2641

Db       2101 CTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTC 2160

Qy       2642 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2701

Db       2161 AATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCA 2220

Qy       2702 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2761

Db       2221 TCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCT 2280

Qy       2762 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2821

Db       2281 TCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCA 2340

Qy       2822 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2881

Db       2341 GGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGAT 2400

Qy       2882 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2941

Db       2401 ATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTG 2460

Qy       2942 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 3001

Db       2461 AAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGA 2520

Qy       3002 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 3061

Db       2521 CTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGA 2580

Qy       3062 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 3121

Db       2581 ACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATG 2640

Qy       3122 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 3181

Db       2641 CAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATAC 2700

Qy       3182 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 3241

Db       2701 AATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGG 2760

Qy       3242 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 3301

Db       2761 AGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGG 2820

Qy       3302 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 3361

Db       2821 TCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGT 2880

Qy       3362 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 3421

Db       2881 GGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTA 2940

Qy       3422 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3481

Db       2941 TATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGC 3000

Qy       3482 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3541

Db       3001 CACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCA 3060

Qy       3542 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3601

Db       3061 ATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAAT 3120

Qy       3602 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3661

Db       3121 GGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTA 3180

Qy       3662 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3721

Db       3181 ATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGA 3240

Qy       3722 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3781

Db       3241 AACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTT 3300

Qy       3782 TGCTGAAAAAAAAA 3794

Db       3301 TGCTGAAAAAAAAA 3313
```

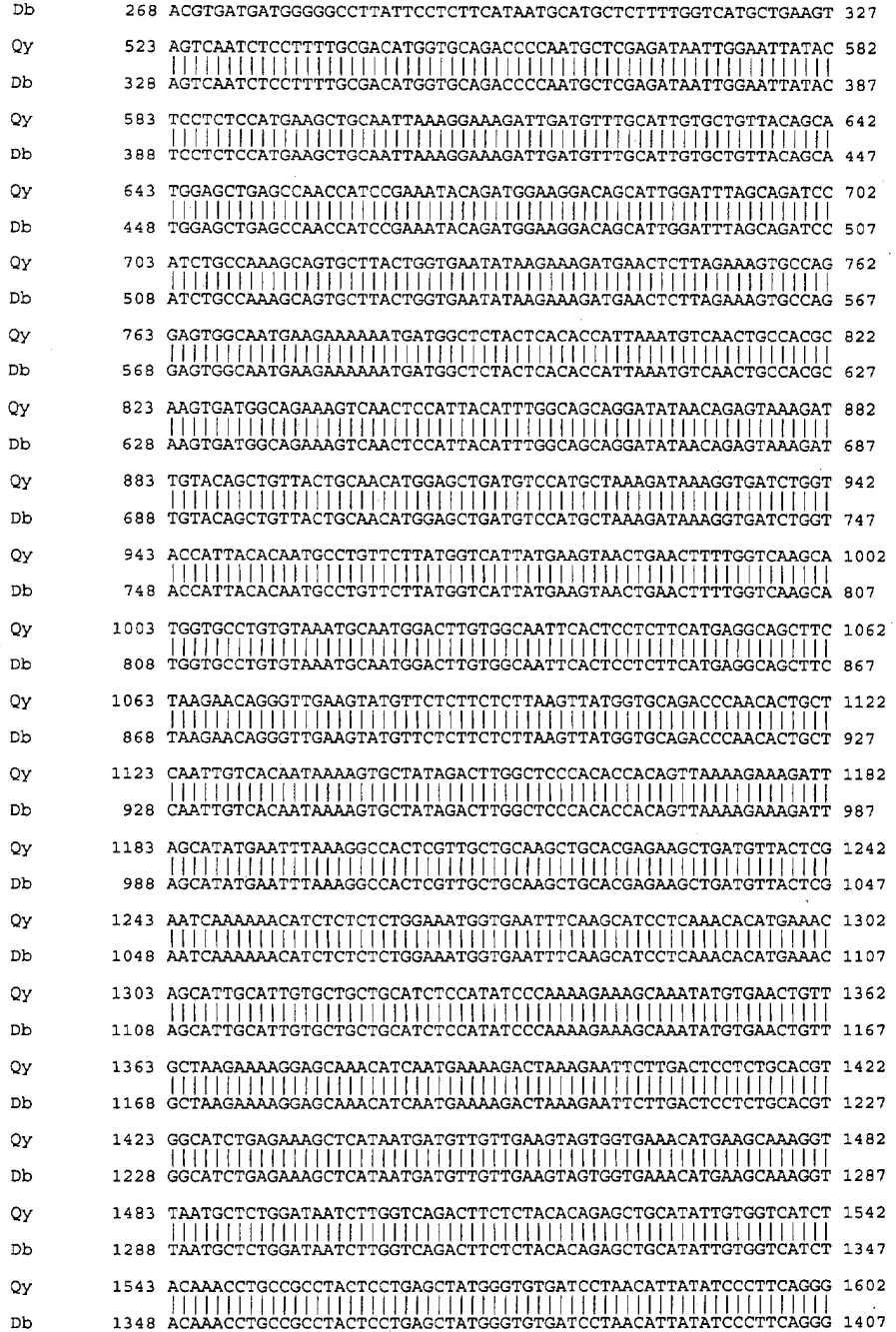
```
RESULT 10
US-10-199-937-1
; Sequence 1, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3508
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(3508)
US-10-199-937-1

  Query Match             86.8%;  Score 3296.2;  DB 14;  Length 3508;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3298;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
```


Db 268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327
Qy 523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db 328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387
Qy 583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db 388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447
Qy 643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db 448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507
Qy 703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db 508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567
Qy 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db 568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627
Qy 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db 688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747
Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1587
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1588 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1647
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1648 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1707
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1708 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1767
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1768 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1827
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 1828 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1887
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1888 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1947
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1948 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2007
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2008 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2067
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2068 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2127
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2128 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2187
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2188 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2247
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2248 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2307
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2308 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2367
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2368 AGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2427
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2428 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2487

```
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2488 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2547

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2548 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2607

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2608 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2667

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2668 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2727

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2728 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2787

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2788 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2847

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2848 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2907

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2908 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2967

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2968 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3027

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3028 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3087

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3088 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3147

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3148 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3207

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3208 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3267

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3268 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3327

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3328 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3387

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3388 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3447

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3448 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3507

Qy       3703 A 3703
             |
Db       3508 A 3508
```

```
RESULT 11
US-09-972-115A-1
; Sequence 1, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3999)
;   OTHER INFORMATION:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(586)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (596)..(597)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (621)..(621)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (626)..(626)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (629)..(629)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (638)..(638)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (693)..(694)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (724)..(724)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (730)..(730)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (764)..(764)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (791)..(791)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2148)..(2148)
```

```
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2177)..(2177)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2382)..(2382)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2398)..(2398)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2466)..(2466)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2468)..(2468)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2493)..(2493)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2536)..(2860)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (3163)..(3163)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (3488)..(3499)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
US-09-972-115A-1

  Query Match            78.2%;  Score 2969.6;  DB 10;  Length 4493;
  Best Local Similarity  88.6%;  Pred. No. 0;
  Matches 3028;  Conservative    1;  Mismatches  363;  Indels   27;  Gaps    2;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||| || | || ||||||||||||| |||||||||| | ||||| |
Db        672 AGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGCTTCAAAATGGTGCAAATNTCCAANC 731

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              || | || ||||||||||||||||||||||| |||||||||||||||||||||| ||
Db        732 ACTTTATAATGGGGGCCTTATTCCTCTTCATANTGCATGCTCTTTTGGTCATGCTAAAAN 791

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 TATCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 851

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 911

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 971

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1031

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1032 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1091

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1151

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1211

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1271

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1331

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1391

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1451

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1452 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1511

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1512 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1571

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1572 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1631

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1632 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1691

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1692 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1751

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1752 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1811

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1812 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1871

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1872 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1931

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1932 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1991

Qy       1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1992 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2051

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2052 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2111

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db       2112 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGNGGCCTTGTACCTTTGCACAATGC 2171

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

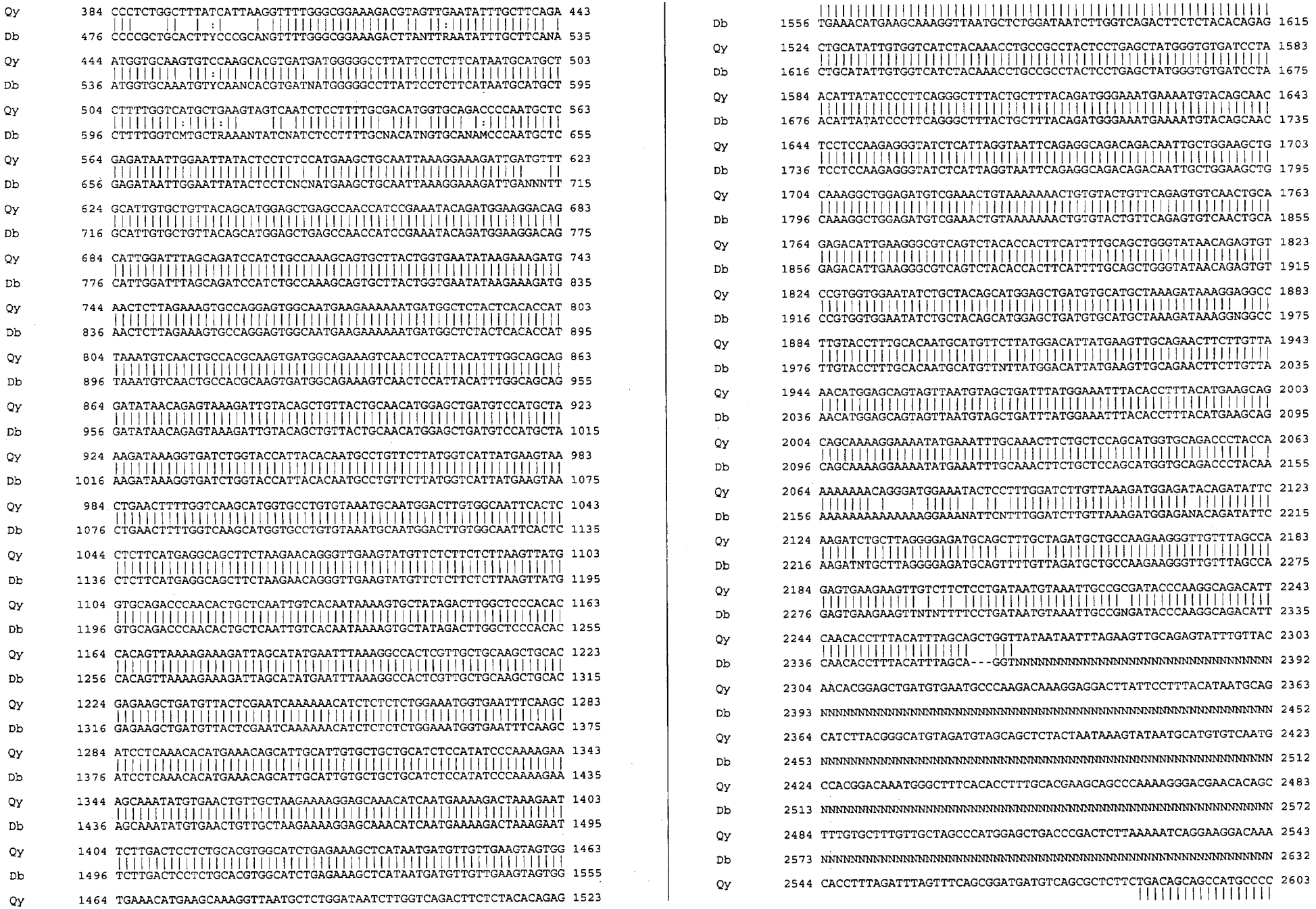
```
Db       2172 ATGTTNTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2231
Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db       2232 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2291
Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db       2292 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2351
Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db       2352 TACTCTTTTGGATCTTGTTAAAGATGGAGANACAGATATTCAAGATNTGCTTAGGGGAGA 2411
Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       2412 TGCAGTTTTGTTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTNTNTTT 2471
Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       2472 TCCTGATAATGTAAATTGCCGNGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2531
Qy       2263 AGCT------GGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGA 2316
Db       2532 AGGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2591
Qy       2317 TGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCA 2376
Db       2592 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2651
Qy       2377 TGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATG 2436
Db       2652 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2711
Qy       2437 GGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTT 2496
Db       2712 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2771
Qy       2497 GCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTT 2556
Db       2772 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2831
Qy       2557 AGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCC 2616
Db       2832 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAGCAGCCATGCCCCCATCTGTTCTGCC 2891
Qy       2617 CTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGC 2676
Db       2892 CTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGC 2951
Qy       2677 TCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTT 2736
Db       2952 TCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTT 3011
Qy       2737 ATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTC 2796
Db       3012 ATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTC 3071
Qy       2797 CAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 2856
Db       3072 CAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAA 3131
Qy       2857 TCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATT 2916
Db       3132 TCTTGGACTTGAGCACCTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGTATT 3191
Qy       2917 AGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCA 2976
Db       3192 AGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGWATCAATGCTTATGGACATAGGCA 3251
Qy       2977 CAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTT 3036
Db       3252 CAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTT 3311
```

```
Qy       3037 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3096
Db       3312 AACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGA 3371
Qy       3097 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3156
Db       3372 GTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCA 3431
Qy       3157 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAA 3216
Db       3432 TGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACANNNN 3491
Qy       3217 ACTATG--------------------GGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3255
Db       3492 NNNNNNNNGAGCCAAGATTCGGCACGAGGAAAGATACACTCACCGGAGAAAAGAAGTTTC 3551
Qy       3256 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3315
Db       3552 TGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAA 3611
Qy       3316 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3375
Db       3612 TGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGC 3671
Qy       3376 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3435
Db       3672 TGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGG 3731
Qy       3436 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3495
Db       3732 AGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCT 3791
Qy       3496 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3555
Db       3792 CTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACA 3851
Qy       3556 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3615
Db       3852 TTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGC 3911
Qy       3616 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3675
Db       3912 TGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGAT 3971
Qy       3676 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 3735
Db       3972 TATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTG 4031
Qy       3736 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       4032 AACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 4090
```

RESULT 12
US-09-972-115A-3
; Sequence 3, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3
; LENGTH: 4297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3801)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (125)..(125)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (293)..(293)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (328)..(328)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (341)..(341)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (378)..(378)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (415)..(415)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (459)..(459)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (496)..(496)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (517)..(517)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (534)..(534)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (553)..(553)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (562)..(562)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (614)..(614)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (619)..(619)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (632)..(632)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (637)..(637)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (643)..(643)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (680)..(680)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (682)..(682)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (711)..(713)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (1971)..(1971)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2000)..(2000)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2175)..(2175)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2180)..(2180)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2205)..(2205)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2221)..(2221)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2289)..(2289)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2316)..(2316)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2359)..(2674)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2977)..(2977)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2291)..(2291)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
US-09-972-115A-3

Query Match 78.2%; Score 2968.2; DB 10; Length 4297;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 3020; Conservative 7; Mismatches 381; Indels 12; Gaps 2;



```
Db       2633 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAGCAGCCATGCCCC 2692
Qy       2604 CATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAG 2663
Db       2693 CATCTGTTCTGCCCTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAG 2752
Qy       2664 CCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCA 2723
Db       2753 CCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCA 2812
Qy       2724 GTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAA 2783
Db       2813 GTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAA 2872
Qy       2784 CAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTC 2843
Db       2873 CAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTC 2932
Qy       2844 AATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCA 2903
Db       2933 AATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTNAGAGAGAACAGATCA 2992
Qy       2904 CTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTT 2963
Db       2993 CTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGWATCAATGCTT 3052
Qy       2964 ATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTC 3023
Db       3053 ATGGACATAGGCACAAACTAATTAAAAGTTTCGAGAGACTTATCTCCGGACAACAAGGTC 3112
Qy       3024 TTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTC 3083
Db       3113 TTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTC 3172
Qy       3084 CTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACA 3143
Db       3173 CTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACA 3232
Qy       3144 GAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGG 3203
Db       3233 GAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGG 3292
Qy       3204 TTTGTAAC---------AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3254
Db       3293 TTTGTAACAGAGCCAAGATTCGGCACGAGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3352
Qy       3255 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3314
Db       3353 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3412
Qy       3315 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3374
Db       3413 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3472
Qy       3375 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3434
Db       3473 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3532
Qy       3435 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3494
Db       3533 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3592
Qy       3495 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 3554
Db       3593 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 3652
Qy       3555 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3614
Db       3653 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3712
Qy       3615 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3674
Db       3713 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3772
```

```
Qy       3675 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3734
Db       3773 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3832
Qy       3735 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       3833 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3892
```

RESULT 13
US-10-199-937-92
; Sequence 92, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
; LENGTH: 2971
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-937-92

Query Match 64.3%; Score 2439.8; DB 14; Length 2971;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1352 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1411
Db          1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60
Qy       1412 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1471
Db         61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120
Qy       1472 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1531
Db        121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180
Qy       1532 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1591
Db        181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240
Qy       1592 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1651
Db        241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300
Qy       1652 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1711
Db        301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360
Qy       1712 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1771
Db        361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420
Qy       1772 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1831
Db        421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480
Qy       1832 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1891
```

```
Db        481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540
Qy       1892 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1951
Db        541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600
Qy       1952 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
Db        601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660
Qy       2012 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2071
Db        661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720
Qy       2072 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2131
Db        721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780
Qy       2132 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2191
Db        781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840
Qy       2192 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2251
Db        841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900
Qy       2252 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2311
Db        901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960
Qy       2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2371
Db        961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020
Qy       2372 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2431
Db       1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080
Qy       2432 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2491
Db       1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140
Qy       2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2551
Db       1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200
Qy       2552 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2611
Db       1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260
Qy       2612 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2671
Db       1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320
Qy       2672 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2731
Db       1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380
Qy       2732 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2791
Db       1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440
Qy       2792 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2851
Db       1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500
Qy       2852 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2911
Db       1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560
Qy       2912 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2971
Db       1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2972 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3031
Db       1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680
Qy       3032 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3091
Db       1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740
Qy       3092 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3151
Db       1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800
Qy       3152 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3211
Db       1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860
Qy       3212 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3271
Db       1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920
Qy       3272 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3331
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980
Qy       3332 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3391
Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040
Qy       3392 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451
Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100
Qy       3452 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3511
Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160
Qy       3512 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3571
Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220
Qy       3572 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3631
Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280
Qy       3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3691
Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340
Qy       3692 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3751
Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400
Qy       3752 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443
```

RESULT 14
US-10-199-937-88
; Sequence 88, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28

```
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
;   LENGTH: 3353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2352)
US-10-199-937-88

  Query Match             64.3%;  Score 2439.8;  DB 14;  Length 3353;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2441;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1352 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 1411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 TGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACT 60

Qy       1412 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 1471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACAT 120

Qy       1472 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 1531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 GAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATAT 180

Qy       1532 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 1591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 TGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATA 240

Qy       1592 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 1651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 TCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAA 300

Qy       1652 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 1711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 GAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCT 360

Qy       1712 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 1771
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 GGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATT 420

Qy       1772 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 1831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 GAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTG 480

Qy       1832 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 1891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 GAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCT 540

Qy       1892 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 1951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           541 TTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGA 600

Qy       1952 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 2011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           601 GCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAA 660

Qy       2012 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAC 2071
             |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db           661 GGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAC 720

Qy       2072 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 2131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           721 AGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTG 780

Qy       2132 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 2191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           781 CTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAG 840

Qy       2192 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 2251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           841 AAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCT 900

Qy       2252 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 2311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           901 TTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGA 960

Qy       2312 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 2371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           961 GCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTAC 1020

Qy       2372 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 2431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1021 GGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGAC 1080

Qy       2432 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 2491
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1081 AAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCT 1140

Qy       2492 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 2551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1141 TTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTA 1200

Qy       2552 GATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 2611
             |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db           1201 GATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCT 1260

Qy       2612 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 2671
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1261 CTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCA 1320

Qy       2672 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 2731
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1321 GATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGAC 1380

Qy       2732 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 2791
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1381 AACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGT 1440

Qy       2792 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 2851
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1441 GCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTA 1500

Qy       2852 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 2911
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1501 AGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGAT 1560

Qy       2912 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 2971
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1561 GTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACAT 1620

Qy       2972 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 3031
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1621 AGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCA 1680

Qy       3032 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 3091
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1681 TATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGAT 1740

Qy       3092 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 3151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1741 AAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGA 1800

Qy       3152 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 3211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1801 GGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAAC 1860

Qy       3212 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 3271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1861 AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAAC 1920

Qy       3272 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 3331
```

```
Db       1921 CATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAA 1980

Qy       3332 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 3391

Db       1981 GGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCT 2040

Qy       3392 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 3451

Db       2041 GAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCA 2100

Qy       3452 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 3511

Db       2101 GTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACC 2160

Qy       3512 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 3571

Db       2161 TTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCAT 2220

Qy       3572 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 3631

Db       2221 CACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTAC 2280

Qy       3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 3691

Db       2281 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGT 2340

Qy       3692 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 3751

Db       2341 ATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAA 2400

Qy       3752 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794

Db       2401 AGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 2443


RESULT 15
US-10-199-937-90
; Sequence 90, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
;   LENGTH: 3799
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (3)..(2270)
US-10-199-937-90

  Query Match           62.1%;  Score 2357.8;  DB 14;  Length 3799;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 2359;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1434 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 1493

Db          1 AAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGG 60

Qy       1494 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 1553

Db         61 ATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCC 120

Qy       1554 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 1613

Db        121 GCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTT 180

Qy       1614 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 1673

Db        181 TACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATT 240

Qy       1674 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 1733

Db        241 CAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAA 300

Qy       1734 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 1793

Db        301 AACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCAC 360

Qy       1794 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 1853

Db        361 TTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAG 420

Qy       1854 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 1913

Db        421 CTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATG 480

Qy       1914 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 1973

Db        481 GACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATT 540

Qy       1974 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 2033

Db        541 TATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAAC 600

Qy       2034 TTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGG 2093

Db        601 TTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGG 660

Qy       2094 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 2153

Db        661 ATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGC 720

Qy       2154 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 2213

Db        721 TAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATG 780

Qy       2214 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 2273

Db        781 TAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATA 840

Qy       2274 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 2333

Db        841 ATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACA 900

Qy       2334 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 2393

Db        901 AAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTC 960

Qy       2394 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 2453

Db        961 TACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGC 1020

Qy       2454 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 2513

Db       1021 ACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTG 1080

Qy       2514 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATG 2573

Db       1081 ACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATG 1140

Qy       2574 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 2633
```

```
Db       1141 TCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTC 1200

Qy       2634 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 2693
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCAT 1260

Qy       2694 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 2753
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAG 1320

Qy       2754 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 2813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGG 1380

Qy       2814 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 2873
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACC 1440

Qy       2874 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 2933
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACA 1500

Qy       2934 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 2993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAG 1560

Qy       2994 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 3053
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTG 1620

Qy       3054 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 3113
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGG 1680

Qy       3114 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 3173
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCA 1740

Qy       3174 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 3233
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACA 1800

Qy       3234 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 3293
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTAT 1860

Qy       3294 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 3353
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGT 1920

Qy       3354 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 3413
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 ACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATC 1980

Qy       3414 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 3473
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT 2040

Qy       3474 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 3533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 ACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGT 2100

Qy       3534 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 3593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCA 2160

Qy       3594 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 3653
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTG 2220

Qy       3654 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 3713
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 AGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTA 2280
```

```
Qy       3714 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 3773
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 TTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTT 2340

Qy       3774 TACTCCTTTGCTGAAAAAAAA 3794
             ||||||||||||||||||||
Db       2341 TACTCCTTTGCTGAAAAAAAA 2361
```

RESULT 16
US-10-199-937-3
; Sequence 3, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 3984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3981)
US-10-199-937-3

Query Match 47.8%; Score 1813.4; DB 14; Length 3984;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
              || | |  || |   |  |||||||||  |||| || || || ||| | || || |||
Db        652 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 711

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
              |  ||||| | |||||| || |||||||||| || || || || |||||||||| ||
Db        712 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 771

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
              |||||||| |||||||| || || | ||| | ||| | || || || || || |||||
Db        772 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 831

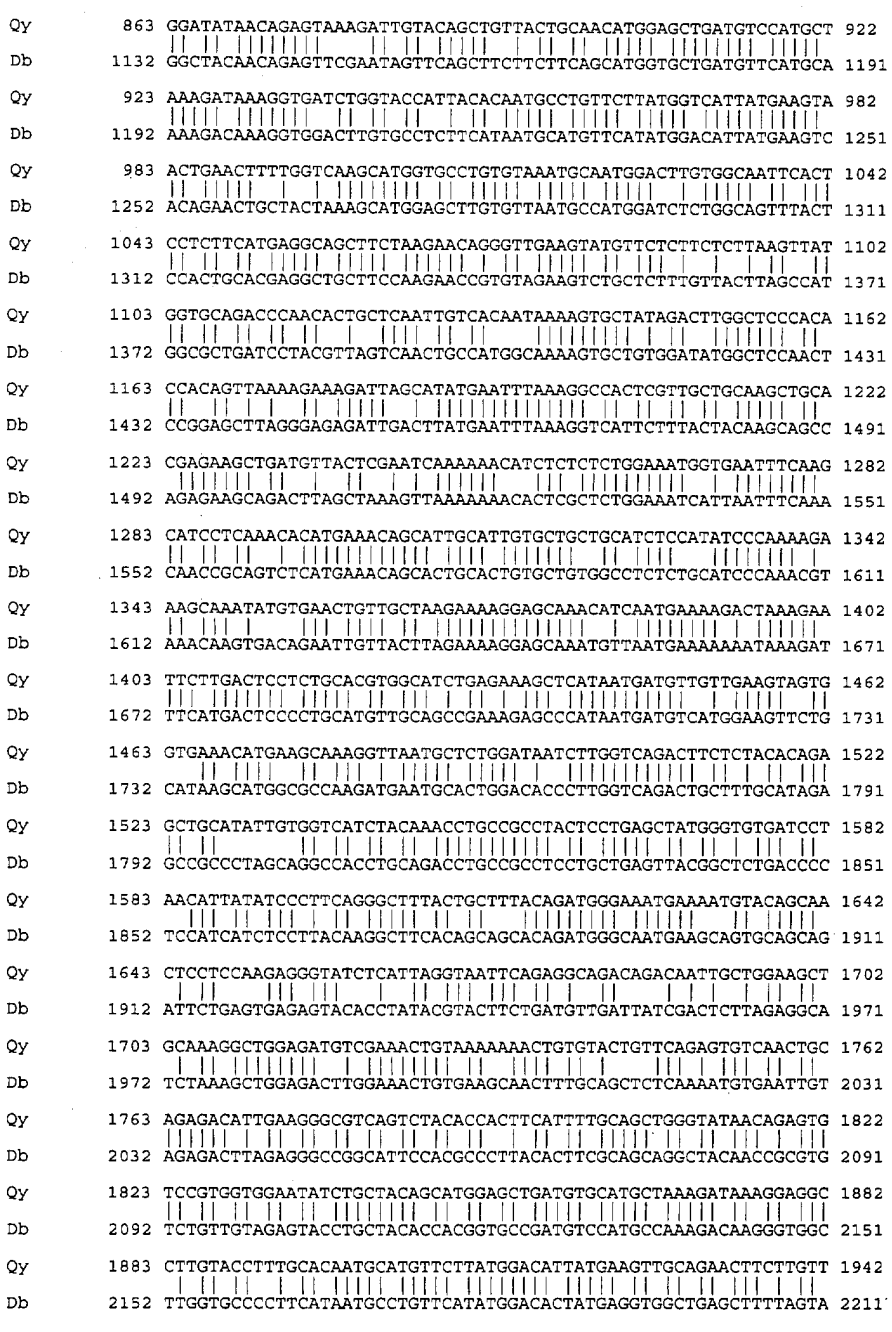
Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
              | ||||| ||||| ||||| ||||| |||||||||| |||||||| ||||| ||||||
Db        832 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 891

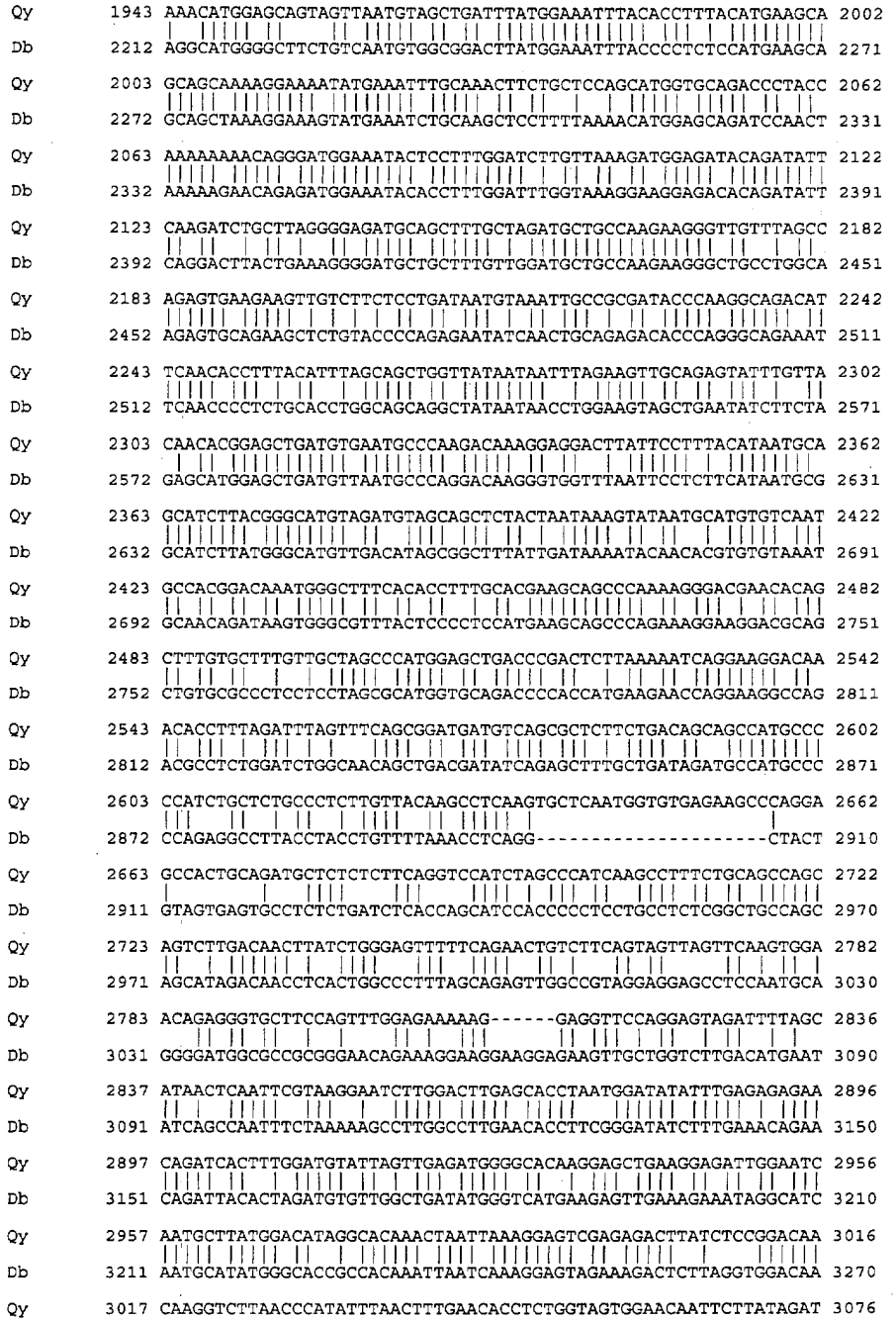
Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
              |||||||||||| | ||||| |||||||| |||| ||| || || || ||||| |  ||
Db        892 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 951

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
              ||  ||||  | |||||||| || || ||||| || ||||| |||||||| ||||||||
Db        952 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1011

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
              |||||| |||||  ||| |||||||| |||||||||| | |||||| |||| || ||
Db        1012 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1071

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
              ||||||| || ||||| |||||||||| ||||||| |||| |||||| | |||||
Db        1072 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1131
```


Qy 863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db 1132 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1191
Qy 923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db 1192 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1251
Qy 983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db 1252 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1311
Qy 1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db 1312 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1371
Qy 1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db 1372 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1431
Qy 1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db 1432 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1491
Qy 1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db 1492 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1551
Qy 1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db 1552 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1611
Qy 1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db 1612 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1671
Qy 1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db 1672 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1731
Qy 1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db 1732 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1791
Qy 1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
Db 1792 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1851
Qy 1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
Db 1852 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1911
Qy 1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
Db 1912 ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCA 1971
Qy 1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
Db 1972 TCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGT 2031
Qy 1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
Db 2032 AGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTG 2091
Qy 1823 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1882
Db 2092 TCTGTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAGACAAGGGTGGC 2151
Qy 1883 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1942
Db 2152 TTGGTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTGAGCTTTTAGTA 2211


Qy 1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2002
Db 2212 AGGCATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTCTCCATGAAGCA 2271
Qy 2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2062
Db 2272 GCAGCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAGCAGATCCAACT 2331
Qy 2063 AAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATT 2122
Db 2332 AAAAAGAACAGAGATGGAAATACACCTTTGGATTTGGTAAAGGAAGGAGACACAGATATT 2391
Qy 2123 CAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCC 2182
Db 2392 CAGGACTTACTGAAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGGGCTGCCTGGCA 2451
Qy 2183 AGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACAT 2242
Db 2452 AGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAAT 2511
Qy 2243 TCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTA 2302
Db 2512 TCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTTCTA 2571
Qy 2303 CAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCA 2362
Db 2572 GAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCTCTTCATAATGCG 2631
Qy 2363 GCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAAT 2422
Db 2632 GCATCTTATGGGCATGTTGACATAGCGGCTTTATTGATAAAATACAACACGTGTGTAAAT 2691
Qy 2423 GCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAG 2482
Db 2692 GCAACAGATAAGTGGGCGTTTACTCCCCTCCATGAAGCAGCCCAGAAAGGAAGGACGCAG 2751
Qy 2483 CTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAA 2542
Db 2752 CTGTGCGCCCTCCTCCTAGCGCATGGTGCAGACCCCACCATGAAGAACCAGGAAGGCCAG 2811
Qy 2543 ACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCC 2602
Db 2812 ACGCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCCC 2871
Qy 2603 CCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGA 2662
Db 2872 CCAGAGGCCTTACCTACCTGTTTTAAACCTCAGG----------CTACT 2910
Qy 2663 GCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGC 2722
Db 2911 GTAGTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGC 2970
Qy 2723 AGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGA 2782
Db 2971 AGCATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCA 3030
Qy 2783 ACAGAGGGTGCTTCCAGTTTGGAGAAAAAG------GAGGTTCCAGGAGTAGATTTTAGC 2836
Db 3031 GGGGATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAAT 3090
Qy 2837 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 2896
Db 3091 ATCAGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAA 3150
Qy 2897 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 2956
Db 3151 CAGATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATC 3210
Qy 2957 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 3016
Db 3211 AATGCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAA 3270
Qy 3017 CAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 3076

```
Db       3271 CAAGGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGAT 3330

Qy       3077 CTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGA 3136

Db       3331 CTTGCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGA 3390

Qy       3137 GAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATT 3196

Db       3391 GAACACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATT 3450

Qy       3197 CAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCT 3256

Db       3451 CAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCT 3510

Qy       3257 GAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAAT 3316

Db       3511 GAGGAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAAT 3570

Qy       3317 GCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCT 3376

Db       3571 GCCATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCC 3630

Qy       3377 GGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA 3436

Db       3631 GGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGA 3690

Qy       3437 GGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTC 3496

Db       3691 GGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTC 3750

Qy       3497 TTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACAT 3556

Db       3751 TTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCAC 3810

Qy       3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616

Db       3811 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3870

Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676

Db       3871 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3930

Qy       3677 ATGAGGCCTGAAG 3689

Db       3931 ATGAAGCCAGAAG 3943


RESULT 17
US-09-841-835-1
; Sequence 1, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4134 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-841-835-1

  Query Match             47.8%;  Score 1813.4;  DB 9;  Length 4134;
  Best Local Similarity   72.3%;  Pred. No. 0;
  Matches 2395;  Conservative    0;  Mismatches  891;  Indels   27;  Gaps    2;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682

Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742

Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802

Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862

Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922

Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982

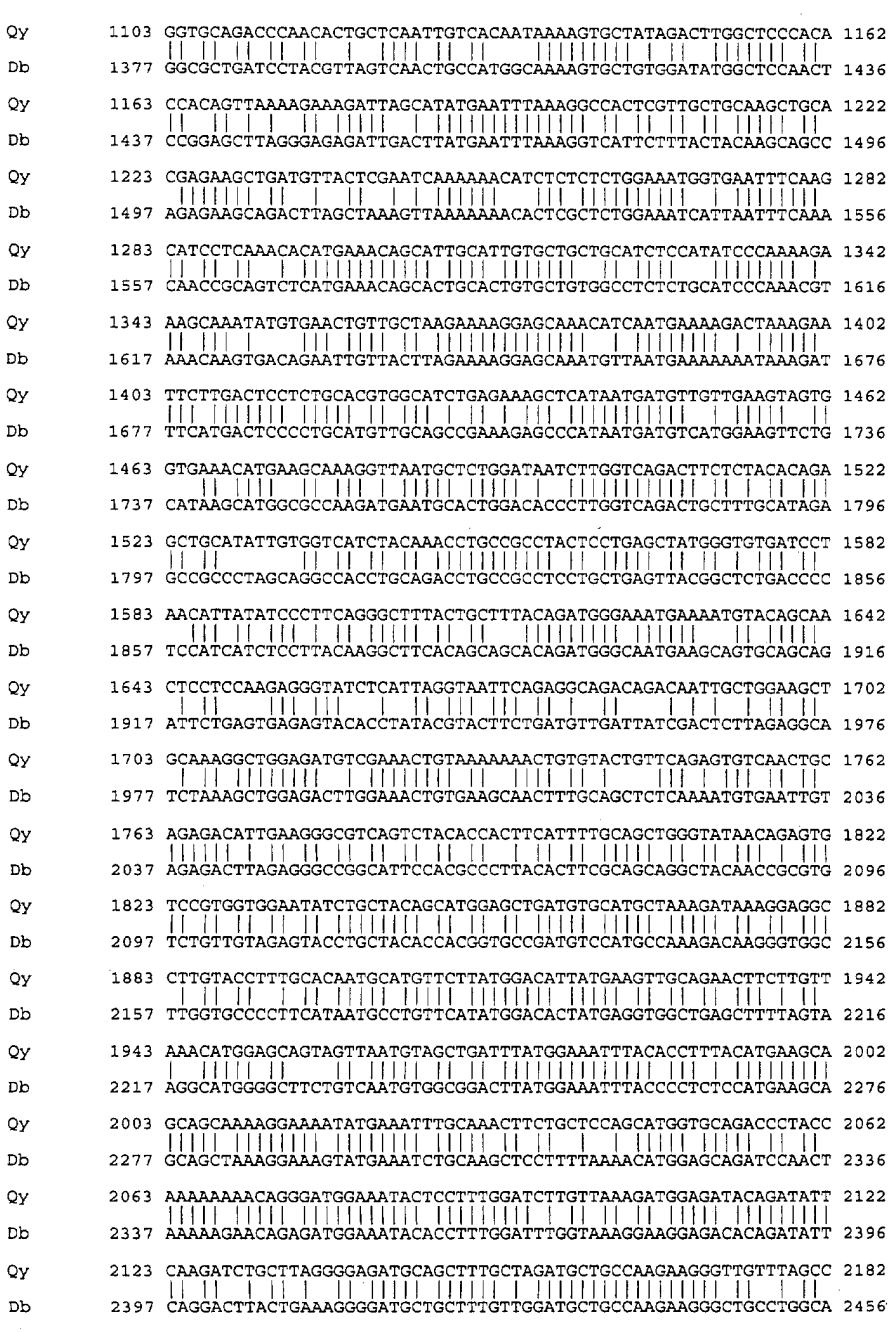
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

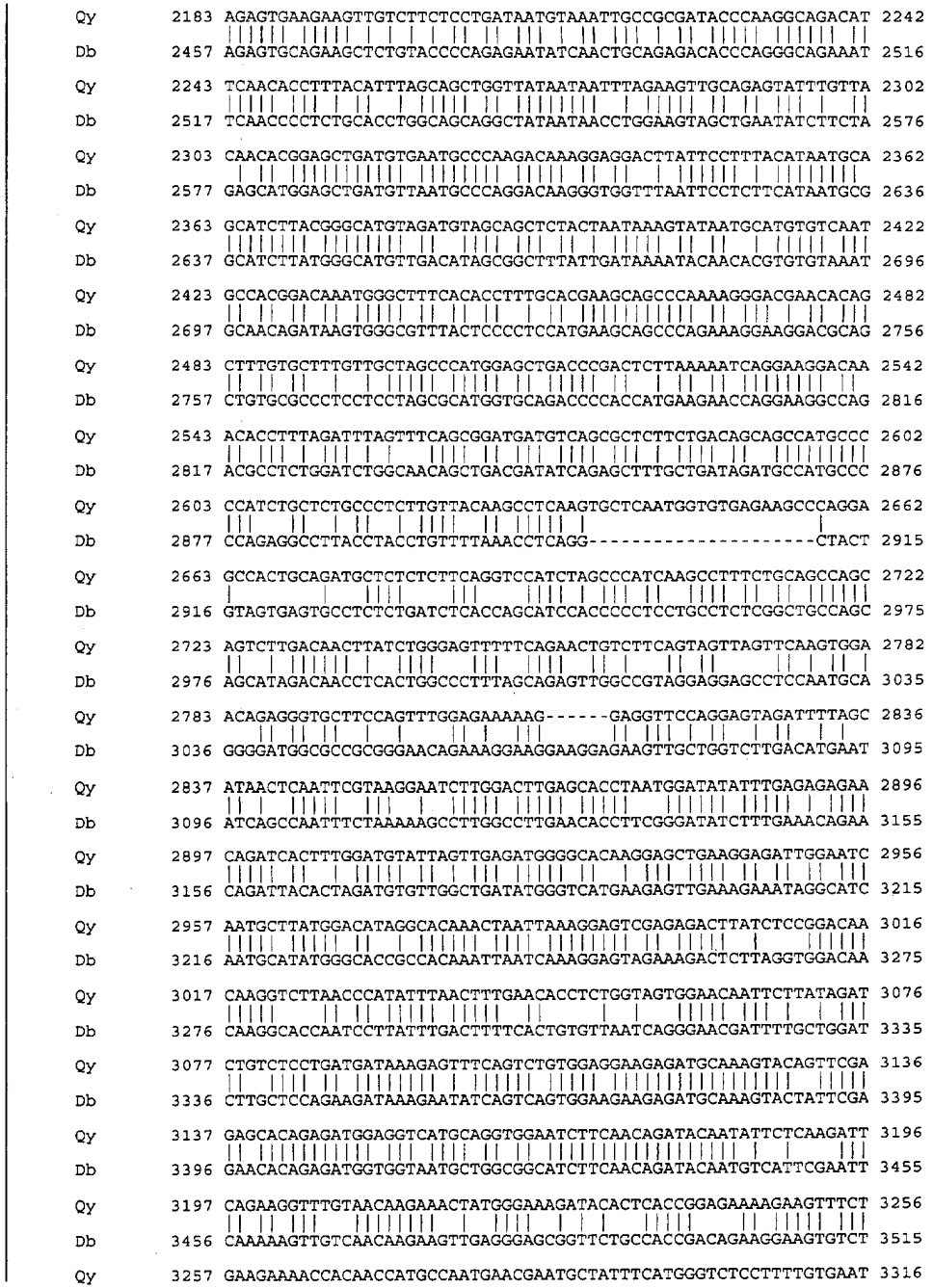
Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042

Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102

Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376
```





```
Db       3516 GAGGAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAAT 3575

Qy       3317 GCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCT 3376

Db       3576 GCCATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCC 3635

Qy       3377 GGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA 3436

Db       3636 GGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGA 3695

Qy       3437 GGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTC 3496

Db       3696 GGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTC 3755

Qy       3497 TTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACAT 3556

Db       3756 TTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCAC 3815

Qy       3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616

Db       3816 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3875

Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676

Db       3876 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3935

Qy       3677 ATGAGGCCTGAAG 3689

Db       3936 ATGAAGCCAGAAG 3948


RESULT 18
US-09-972-115A-7
; Sequence 7, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7 .
;   LENGTH: 4134
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-7

  Query Match             47.8%;  Score 1813.4;  DB 10;  Length 4134;
  Best Local Similarity   72.3%;  Pred. No. 0;
  Matches 2395;  Conservative    0;  Mismatches  891;  Indels   27;  Gaps    2;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682

Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742

Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802

Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862

Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922

Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982

Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042

Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102

Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162

Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222

Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282

Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342

Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402

Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462

Db       1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736

Qy       1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522

Db       1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796

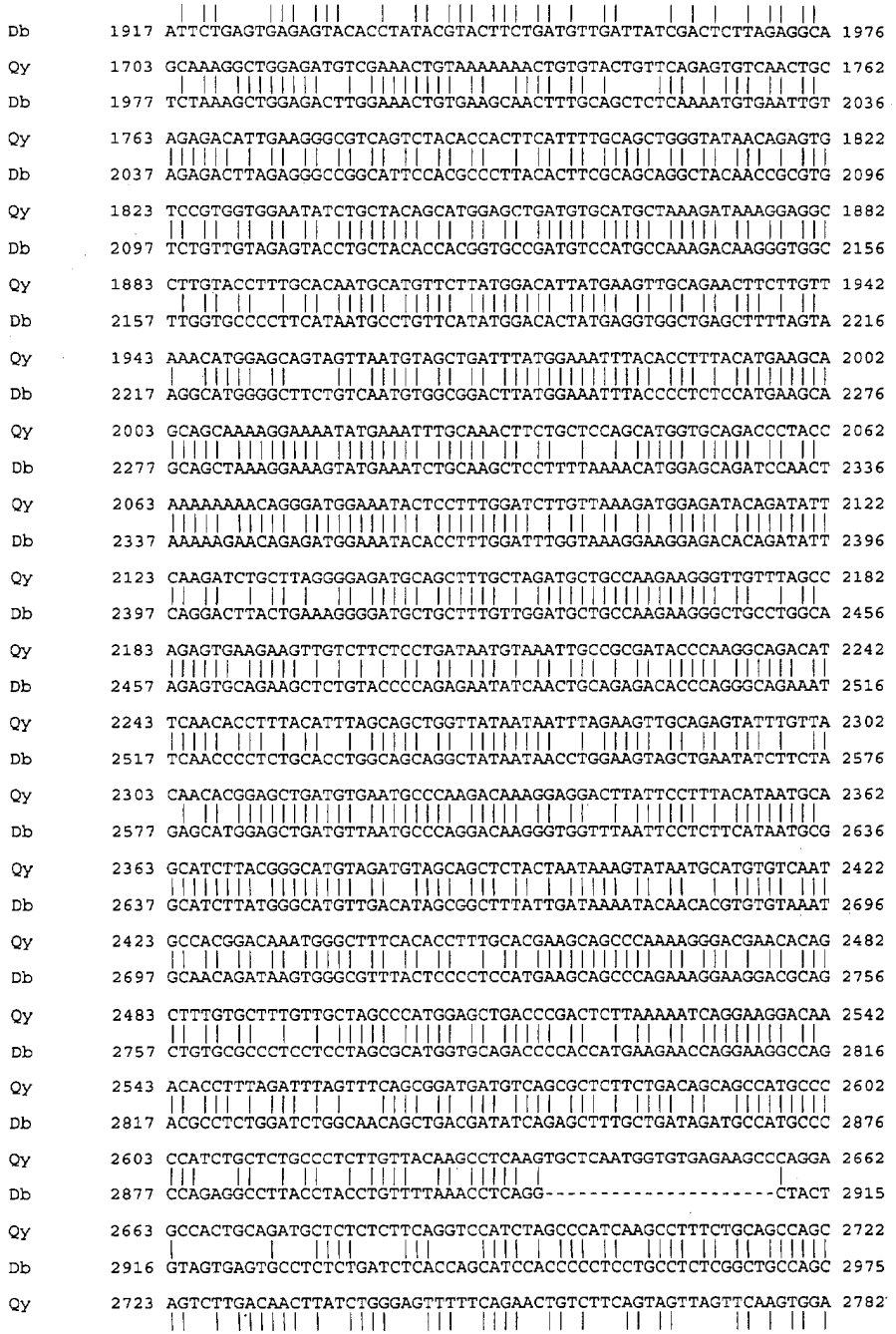
Qy       1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582

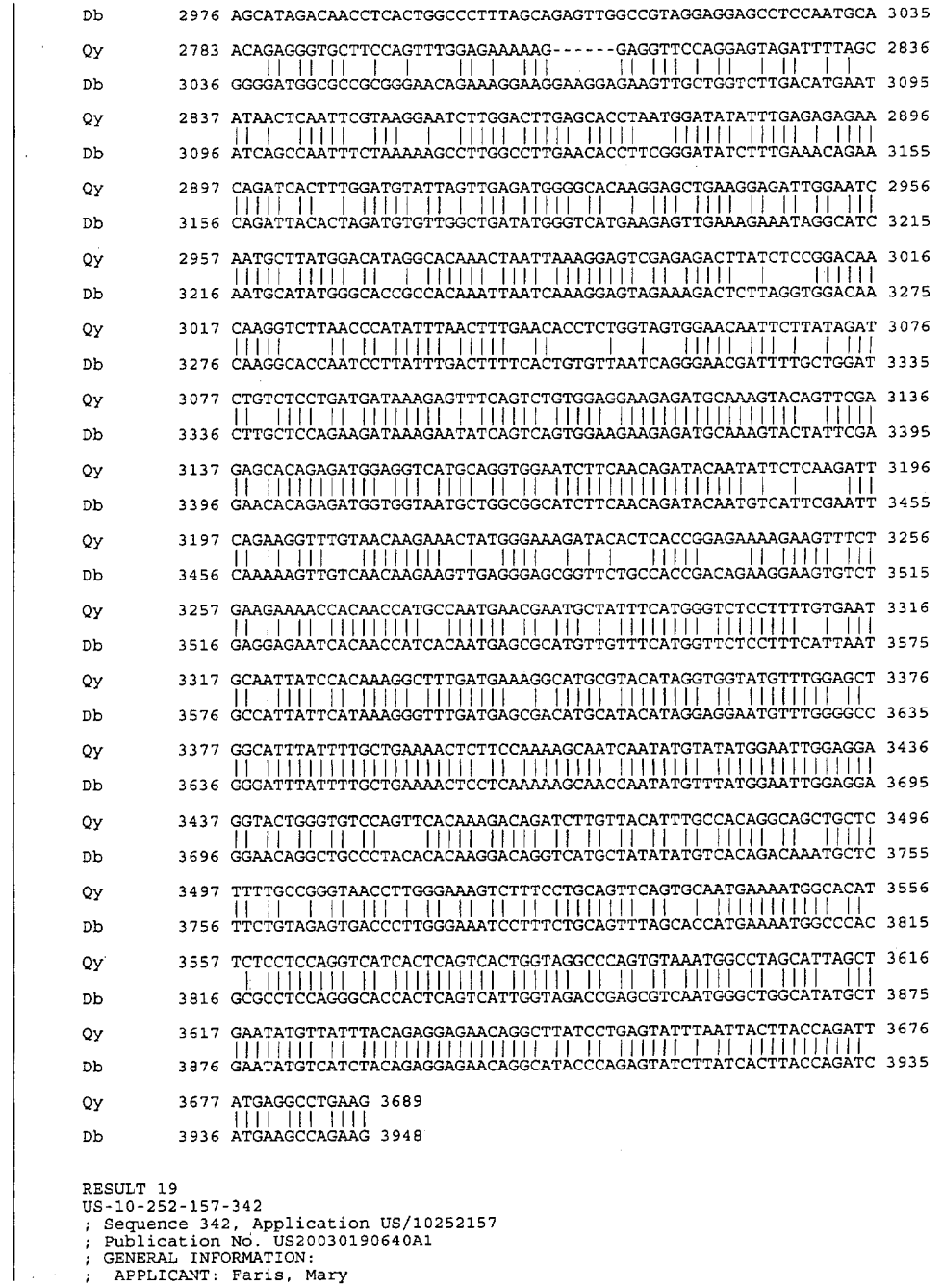
Db       1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856

Qy       1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642

Db       1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916

Qy       1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
```





RESULT 19
US-10-252-157-342
; Sequence 342, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary

```
;  APPLICANT:  Pearson, Cecelia I.
;  TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
;  FILE REFERENCE: PA-0027-1 US
;  CURRENT APPLICATION NUMBER: US/10/252,157
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: 60/295,048
;  PRIOR FILING DATE: 2001-05-31
;  NUMBER OF SEQ ID NOS: 501
;  SOFTWARE: PERL Program
; SEQ ID NO 342
;   LENGTH: 4435
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030190640A1 474266.2
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 4369, 4374
;   OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-342

  Query Match             40.4%;  Score 1533.8;  DB 14;  Length 4435;
  Best Local Similarity   97.1%;  Pred. No. 0;
  Matches 1562;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       2186 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2245
             || || |||| |   | | ||| | | |     | | |||| |   | |
Db        289 GTTAACAAGTGGAAGTATTATGACATAGGACTAAAAAAGAAAAAACAAGTAAAGTAATTT 348

Qy       2246 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2305
                ||   |  |||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TACTTCTTTTCTAAATTAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 408

Qy       2306 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 468

Qy       2366 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 528

Qy       2426 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 588

Qy       2486 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 648

Qy       2546 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 708

Qy       2606 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 768

Qy       2666 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 828

Qy       2726 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 888

Qy       2786 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 2845
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        889 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAA 948

Qy       2846 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 2905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 TTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACT 1008

Qy       2906 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 2965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1009 TTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTAT 1068

Qy       2966 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 3025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1069 GGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTT 1128

Qy       3026 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 3085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1129 AACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCT 1188

Qy       3086 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 3145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1189 GATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGA 1248

Qy       3146 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 3205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1249 GATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTT 1308

Qy       3206 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 3265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1309 TGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAAC 1368

Qy       3266 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 3325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1369 CACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATC 1428

Qy       3326 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 3385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1429 CACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTAT 1488

Qy       3386 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 3445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1489 TTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGG 1548

Qy       3446 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 3505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1549 TGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGG 1608

Qy       3506 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 3565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1609 GTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCA 1668

Qy       3566 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 3625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1669 GGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTT 1728

Qy       3626 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 3685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1729 ATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCT 1788

Qy       3686 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 3745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1789 GAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAAT 1848

Qy       3746 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        1849 CATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 1897

RESULT 20
US-09-841-835-7
; Sequence 7, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Klauber & Jackson
;       STREET:  411 Hackensack Avenue, 4th Floor
;       CITY:  Hackensack
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/841,835
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  09/196,387
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-487-5800
;       TELEFAX:  201-343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4491 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  6..2027
US-09-841-835-7

  Query Match             38.2%;  Score 1449.6;  DB 9;  Length 4491;
  Best Local Similarity   65.3%;  Pred. No. 0;
  Matches 2397;  Conservative    0;  Mismatches  889;  Indels  384;  Gaps    3;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db       1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736

Qy       1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db       1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796

Qy       1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
Db       1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856

Qy       1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
Db       1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916

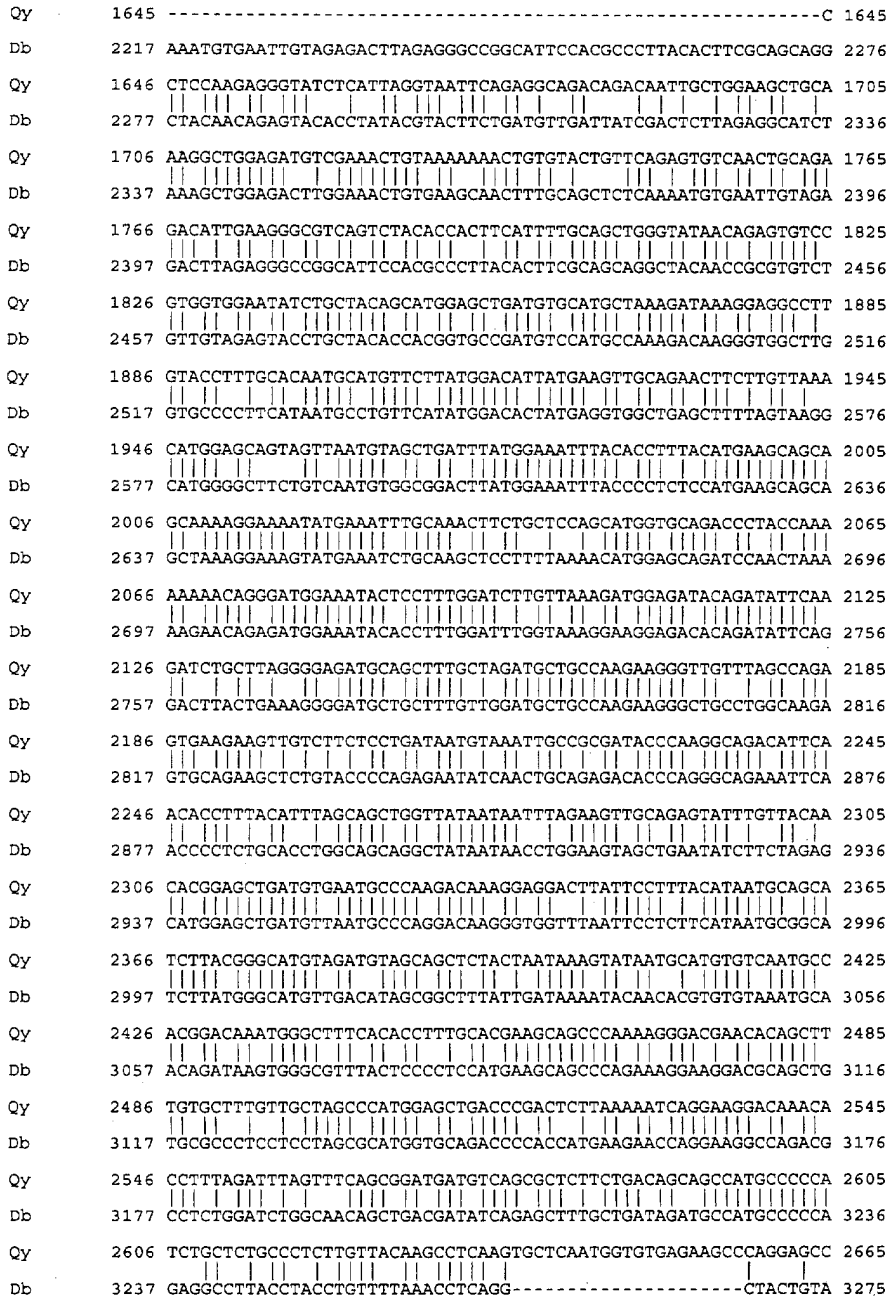
Qy       1643 CT---------------------------------------------------------- 1644
Db       1917 ATTCTGAGTGTGAGTTACGGCTCTGACCCCTCCATCATCTCCTTACAAGGCTTCACAGCA 1976

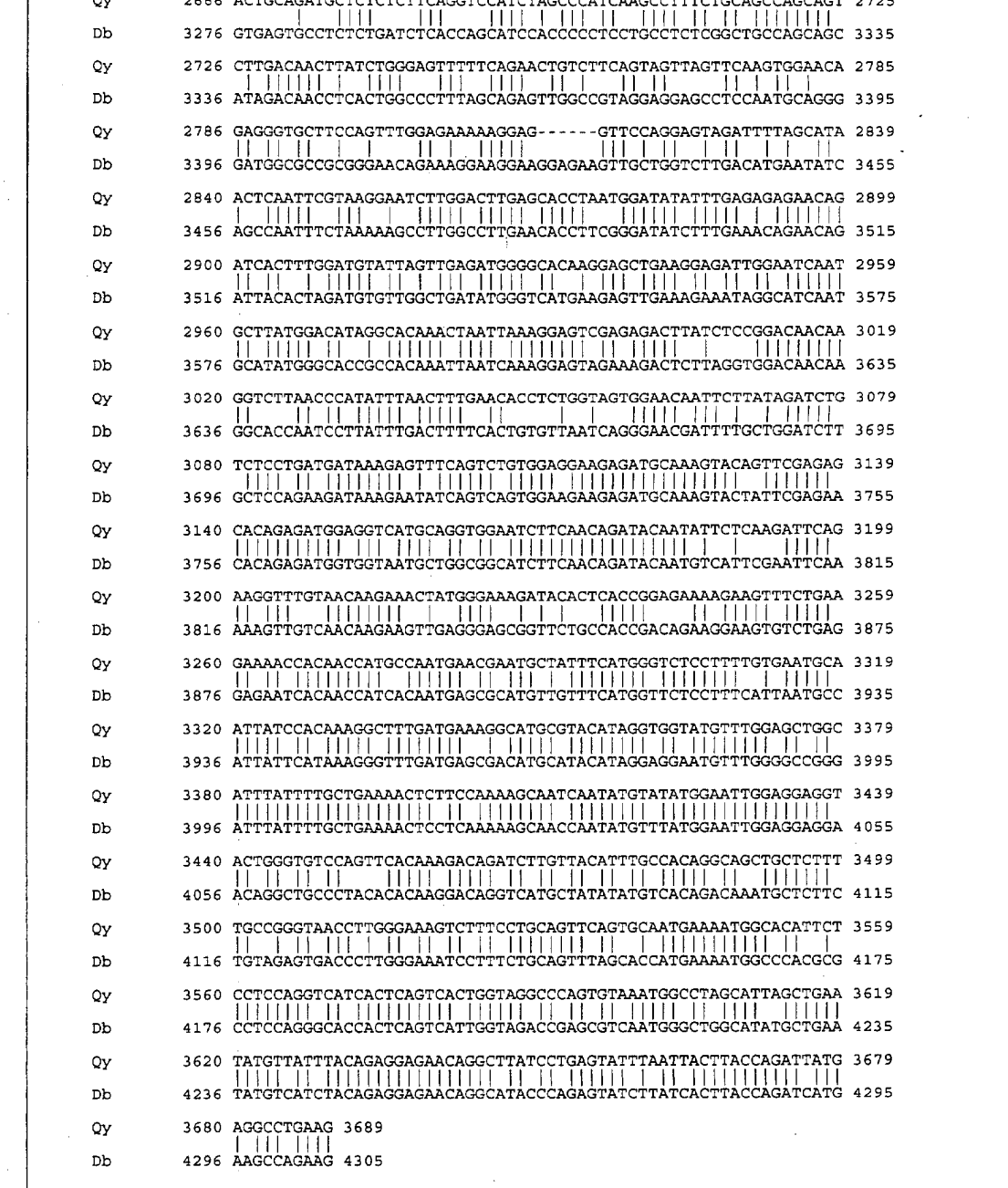
Qy       1645 ---------------------------------------------------------- 1644
Db       1977 GCACAGATGGGCAATGAAGCAGTGCAGCAGATTCTGAGTGGTCATTCGTAGATAGTGATC 2036

Qy       1645 ---------------------------------------------------------- 1644
Db       2037 ATTCTACTTCAGCCTTAATGGTGATCTTGAGACGGGAAGATTTAGAAGGAAATCTATCCA 2096

Qy       1645 ---------------------------------------------------------- 1644
Db       2097 GCATGTCTTCACTGTCAACATGAAGAGTACACCTATACGTACTTCTGATGTTGATTATCG 2156

Qy       1645 ---------------------------------------------------------- 1644
Db       2157 ACTCTTAGAGGCATCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCA 2216
```


Qy 1645 ----------C 1645
Db 2217 AAATGTGAATTGTAGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGG 2276
Qy 1646 CTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCA 1705
Db 2277 CTACAACAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCATCT 2336
Qy 1706 AAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGA 1765
Db 2337 AAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGTAGA 2396
Qy 1766 GACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCC 1825
Db 2397 GACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTGTCT 2456
Qy 1826 GTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTT 1885
Db 2457 GTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAGACAAGGGTGGCTTG 2516
Qy 1886 GTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAA 1945
Db 2517 GTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTGAGCTTTTAGTAAGG 2576
Qy 1946 CATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCA 2005
Db 2577 CATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTCTCCATGAAGCAGCA 2636
Qy 2006 GCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAA 2065
Db 2637 GCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAGCAGATCCAACTAAA 2696
Qy 2066 AAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAA 2125
Db 2697 AAGAACAGAGATGGAAATACACCTTTGGATTTGGTAAAGGAAGGAGACACAGATATTCAG 2756
Qy 2126 GATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGA 2185
Db 2757 GACTTACTGAAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGGGCTGCCTGGCAAGA 2816
Qy 2186 GTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCA 2245
Db 2817 GTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAATTCA 2876
Qy 2246 ACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAA 2305
Db 2877 ACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTTCTAGAG 2936
Qy 2306 CACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCA 2365
Db 2937 CATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCTCTTCATAATGCGGCA 2996
Qy 2366 TCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCC 2425
Db 2997 TCTTATGGGCATGTTGACATAGCGGCTTTATTGATAAAATACAACACGTGTGTAAATGCA 3056
Qy 2426 ACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTT 2485
Db 3057 ACAGATAAGTGGGCGTTTACTCCCCTCCATGAAGCAGCCCAGAAAGGAAGGACGCAGCTG 3116
Qy 2486 TGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACA 2545
Db 3117 TGCGCCCTCCTCCTAGCGCATGGTGCAGACCCCACCATGAAGAACCAGGAAGGCCAGACG 3176
Qy 2546 CCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCA 2605
Db 3177 CCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCCCCCA 3236
Qy 2606 TCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCC 2665
Db 3237 GAGGCCTTACCTACCTGTTTTAAACCTCAGG----------CTACTGTA 3275


Qy 2666 ACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGT 2725
Db 3276 GTGAGTGCCTCTCTGATCTCACCAGCATCCACCCCCTCCTGCCTCTCGGCTGCCAGCAGC 3335
Qy 2726 CTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACA 2785
Db 3336 ATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCAGGG 3395
Qy 2786 GAGGGTGCTTCCAGTTTGGAGAAAAAGGAG------GTTCCAGGAGTAGATTTTAGCATA 2839
Db 3396 GATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAATATC 3455
Qy 2840 ACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAG 2899
Db 3456 AGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAACAG 3515
Qy 2900 ATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAAT 2959
Db 3516 ATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATCAAT 3575
Qy 2960 GCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAA 3019
Db 3576 GCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAACAA 3635
Qy 3020 GGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTG 3079
Db 3636 GGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGATCTT 3695
Qy 3080 TCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAG 3139
Db 3696 GCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGAGAA 3755
Qy 3140 CACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAG 3199
Db 3756 CACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATTCAA 3815
Qy 3200 AAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAA 3259
Db 3816 AAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAG 3875
Qy 3260 GAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCA 3319
Db 3876 GAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCC 3935
Qy 3320 ATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGC 3379
Db 3936 ATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGG 3995
Qy 3380 ATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGT 3439
Db 3996 ATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGA 4055
Qy 3440 ACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTT 3499
Db 4056 ACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTC 4115
Qy 3500 TGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCT 3559
Db 4116 TGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCG 4175
Qy 3560 CCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAA 3619
Db 4176 CCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAA 4235
Qy 3620 TATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATG 3679
Db 4236 TATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATG 4295
Qy 3680 AGGCCTGAAG 3689
Db 4296 AAGCCAGAAG 4305

```
RESULT 21
US-10-199-937-98
; Sequence 98, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
;   LENGTH: 1568
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)..(1567)
US-10-199-937-98

  Query Match           35.8%;  Score 1359.4;  DB 14;  Length 1568;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1360;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        208 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 267

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 327

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 387

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 447

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 507

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 567

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 627

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 687

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 747

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 807

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 867

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 927

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 987

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1047

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1107

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1167

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1227

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1287

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1347

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1407

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1467

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1527

Qy       1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1763
              |||||||||||||||||||||||||||||||||||||||||||
Db       1528 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1568


RESULT 22
US-09-841-835-9
; Sequence 9, Application US/09841835
; Patent No. US20020076795A1
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/841,835
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  09/196,387
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-487-5800
;       TELEFAX:  201-343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4657 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  6..2855
US-09-841-835-9

  Query Match             34.2%;  Score 1297.8;  DB 9;  Length 4657;
  Best Local Similarity   62.9%;  Pred. No. 0;
  Matches 2400;  Conservative    0;  Mismatches  907;  Indels  508;  Gaps    5;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
              || | | .|| |    |  |||||||||  |||| || || || ||| | || || |||
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
              |  |||||  | |||||| || ||||||||||| || || || || ||||||||||| ||
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
              |||||||| |||||||| || || | |||  | ||| | || || || || || |||||
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
              | ||||| ||||| ||||| ||||| |||||||||| |||||||| ||||| |||||
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
              |||||||||||| | ||||| |||||||| |||| ||| || || |||||  |    ||
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
              ||  ||||  | |||||||| || || ||||| || ||||| |||||||| ||||||||
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
              |||||| |||||  ||| |||||||| |||||||||| | |||||| |||| || ||
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
              |||||||  || ||||| |||||||||||  |||||||  ||||| ||||||  | |||||
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
              || || ||||||||    || || |||||  | || || ||||| |||||||| |||||
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
              ||||| |||||||  || || ||  | || |||||  ||||| ||||| ||||||||||
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
              || |||||  |  | |||||||| || |||||  |||||  |||||  | |||||  ||| |||
Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
              || || || |||||  |||||  ||||||  | || |||||  || |||  |  | ||  ||
Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
              || || || ||  ||  |||| || ||  |||||||||  | ||  |||||||  || ||
Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
              ||  || | |  || |||||  | ||||||||||||||  || ||  ||  |||||  ||
Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
              |||||||  ||  |  ||  |  ||||||  ||| ||||||||||  | ||||||||
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
              ||  ||  ||   | ||||||||||||  |||| ||||||  ||  ||||   |||||||| |
Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
              || |||  |    |||  |||| || ||||||||||||||  | |||||||| | ||||||
Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
              ||| |||||| |||||  || ||| | ||  | ||| ||||||||||  | |||||  ||
Db       1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736

Qy       1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
                 || ||||  || ||| | |||||  |||||  | ||||||||||| || | || |||
Db       1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796

Qy       1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
              || ||     || || || || |||||||||| || |||||  || || | |||  ||
Db       1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856

Qy       1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
              |||  || ||| | || |||||  || ||  |||||||||  ||||||   || |||||
Db       1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916

Qy       1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
              | ||    ||| |||    | || ||| ||| ||  ||    | | |  | || ||
Db       1917 ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCA 1976

Qy       1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
              | || |||||||  | |||||||  |||| ||  |    ||| || || ||  ||
Db       1977 TCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGT 2036

Qy       1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
              ||||||  | ||  ||  ||  ||  ||  ||  |  || ||  ||||  || || ||| | |||
Db       2037 AGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTG 2096

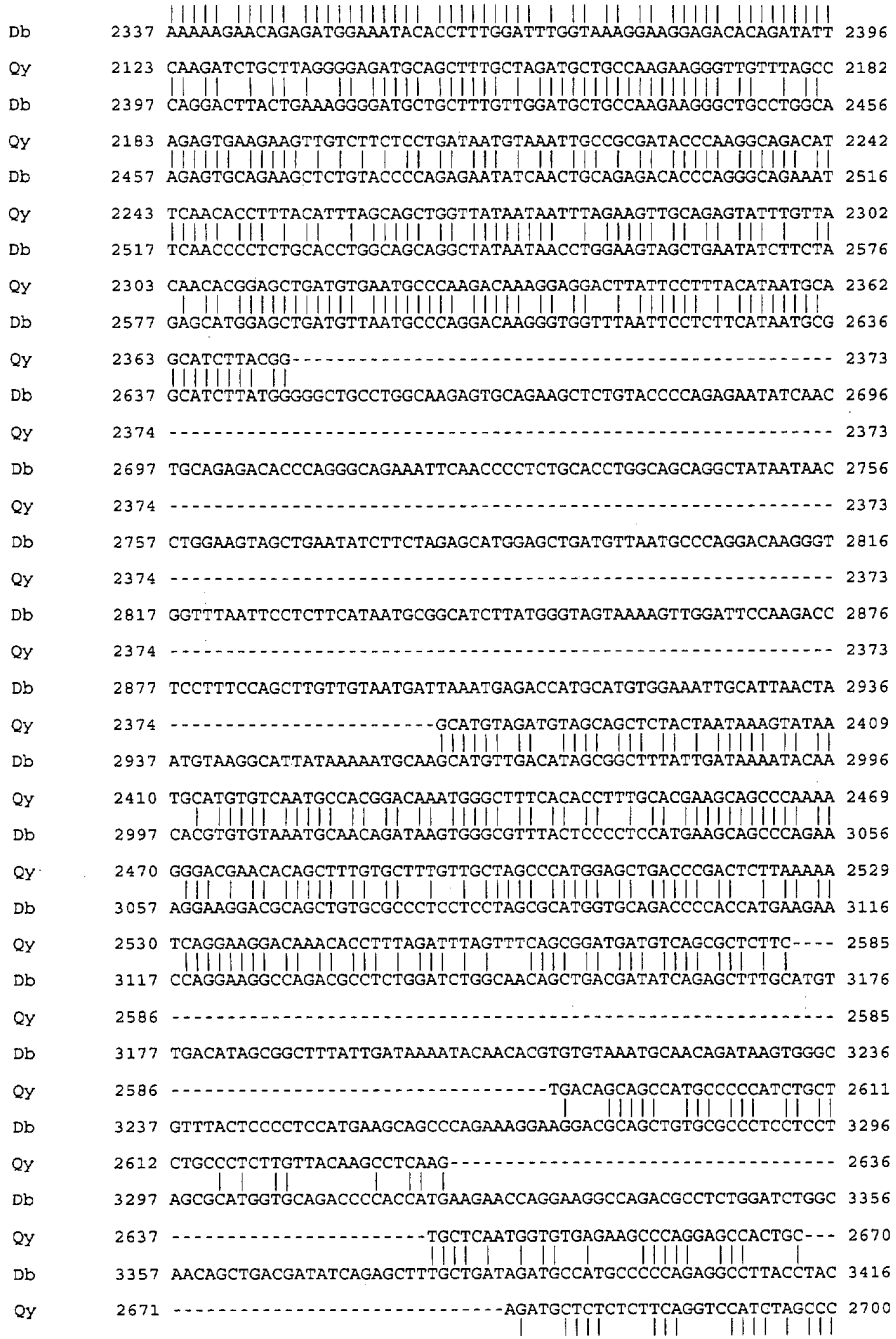
Qy       1823 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1882
              || || || || ||  |||||||| || || || |||||  |||||  |||||  || || |||
Db       2097 TCTGTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAGACAAGGGTGGC 2156

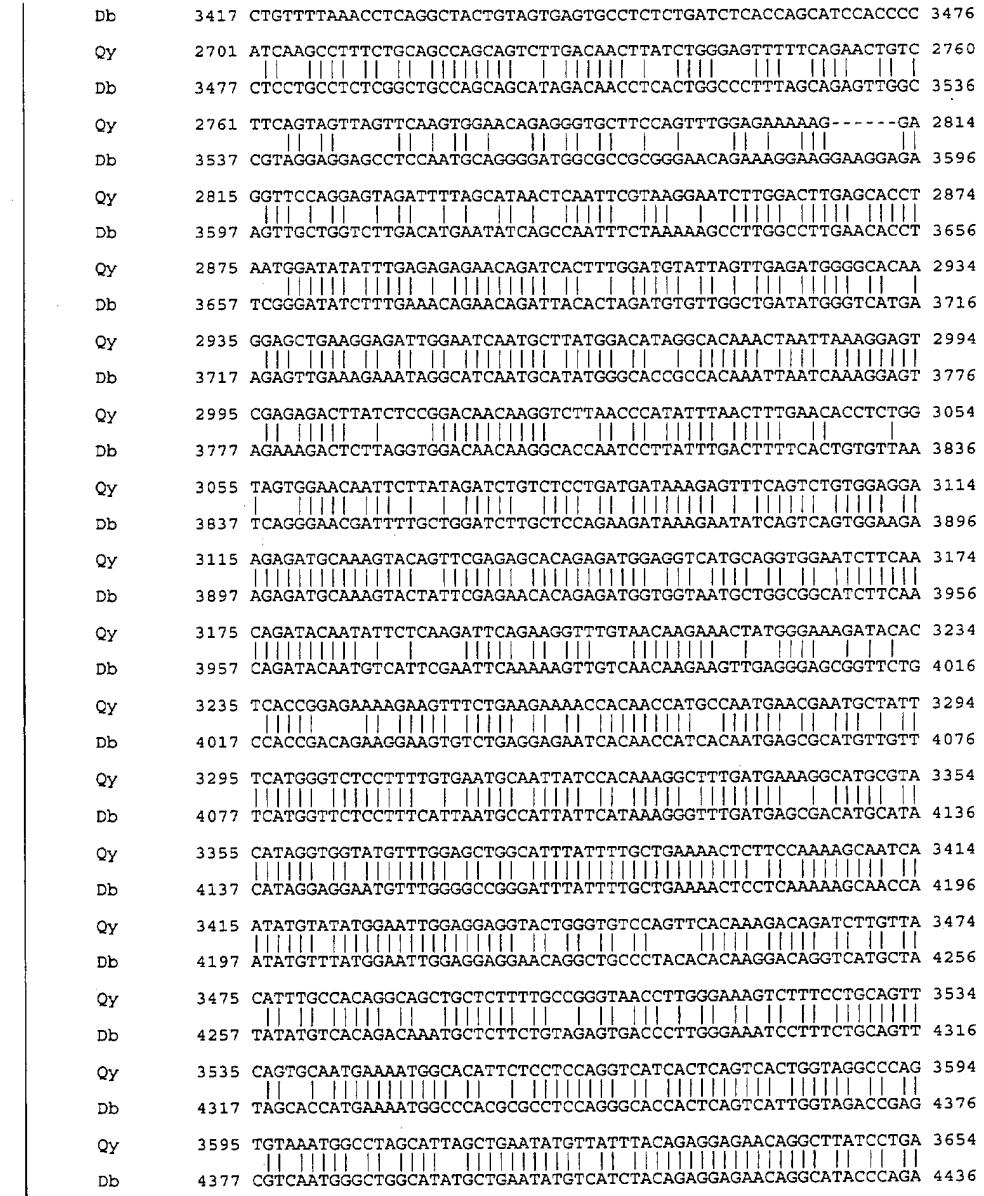
Qy       1883 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1942
              | || ||  | || |||||  |||||  |||||||| || || || |||  | ||
Db       2157 TTGGTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTGAGCTTTTAGTA 2216

Qy       1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2002
              |  |||||  ||    || |||||  || ||  |||||||||||||||  ||| | |||||||||
Db       2217 AGGCATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTCTCCATGAAGCA 2276

Qy       2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2062
              |||| |||||||| |||||  || ||  |  |||||  |||| || || ||||||  || ||
Db       2277 GCAGCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAGCAGATCCAACT 2336

Qy       2063 AAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATT 2122
```





RESULT 23
US-10-276-774-340

```
; Sequence 340, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT:  Tang, Y, Tom et al
;  TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-030
;  CURRENT APPLICATION NUMBER: US/10/276,774
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 2700
;  SOFTWARE: Custom
; SEQ ID NO 340
;   LENGTH: 1466
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-276-774-340

  Query Match            32.7%;  Score 1240.4;  DB 12;  Length 1466;
  Best Local Similarity  99.1%;  Pred. No. 0;
  Matches 1247;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1023 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 1082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTAT 60

Qy       1083 GTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 1142
             ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         61 GTTCTCTTCTCTTAAGTTATGGGGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTG 120

Qy       1143 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 1202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCC 180

Qy       1203 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 1262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTC 240

Qy       1263 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 1322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTG 300

Qy       1323 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 1382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACA 360

Qy       1383 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 1442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATA 420

Qy       1443 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 1502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTG 480

Qy       1503 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 1562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCC 540

Qy       1563 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 1622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGG 600

Qy       1623 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 1682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAG 660

Qy       1683 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTA 1742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTA 720

Qy       1743 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 1802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 CTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTG 780

Qy       1803 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 1862
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGC 840

Qy       1863 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 1922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 ATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATG 900

Qy       1923 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 1982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 AAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAAT 960

Qy       1983 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 2042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 TTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCC 1020

Qy       2043 AGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 2102
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTA 1080

Qy       2103 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 2162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 1140

Qy       2163 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 2222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCC 1200

Qy       2223 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
             ||||||||||||||||||||||||||||||||||||||||||||||||| |  |  |  || ||
Db        1201 GCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGGTAAGTGAATGAAGTT 1258
```

```
RESULT 24
US-10-199-937-177
; Sequence 177, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 177
;   LENGTH: 3308
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Parpla-Tank2b
;   OTHER INFORMATION: Fusion
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3297)
US-10-199-937-177

  Query Match            30.8%;  Score 1168.4;  DB 14;  Length 3308;
  Best Local Similarity  99.9%;  Pred. No. 1.1e-301;
  Matches 1169;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2534 GAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCA 2593
```

```
Db       2101 GAAGGACAAACACCTTTAGATTTAGTTTCAGCAGATGATGTCAGCGCTCTTCTGACAGCA 2160
Qy       2594 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2653
Db       2161 GCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGA 2220
Qy       2654 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2713
Db       2221 AGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCT 2280
Qy       2714 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2773
Db       2281 GCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGT 2340
Qy       2774 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2833
Db       2341 TCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTT 2400
Qy       2834 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2893
Db       2401 AGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGA 2460
Qy       2894 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2953
Db       2461 GAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGA 2520
Qy       2954 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 3013
Db       2521 ATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGA 2580
Qy       3014 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 3073
Db       2581 CAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATA 2640
Qy       3074 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 3133
Db       2641 GATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTT 2700
Qy       3134 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 3193
Db       2701 CGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAG 2760
Qy       3194 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 3253
Db       2761 ATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTT 2820
Qy       3254 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 3313
Db       2821 TCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTG 2880
Qy       3314 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 3373
Db       2881 AATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGA 2940
Qy       3374 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3433
Db       2941 GCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGA 3000
Qy       3434 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3493
Db       3001 GGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTG 3060
Qy       3494 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3553
Db       3061 CTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCA 3120
Qy       3554 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3613
Db       3121 CATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTA 3180
Qy       3614 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3673
Db       3181 GCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAG 3240
Qy       3674 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
Db       3241 ATTATGAGGCCTGAAGGTATGGTCGATGGA 3270
```

RESULT 25
US-10-199-937-159
; Sequence 159, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 159
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (9)..(1094)
US-10-199-937-159

Query Match 28.6%; Score 1084.8; DB 14; Length 1103;
Best Local Similarity 99.8%; Pred. No. 1.3e-279;
Matches 1086; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1037 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 1096
Db          9 TTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTA 68
Qy       1097 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 1156
Db         69 AGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCT 128
Qy       1157 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 1216
Db        129 CCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAA 188
Qy       1217 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 1276
Db        189 GCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAAT 248
Qy       1277 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 1336
Db        249 TTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCC 308
Qy       1337 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 1396
Db        309 AAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACT 368
Qy       1397 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 1456
Db        369 AAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAA 428
Qy       1457 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 1516
Db        429 GTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTA 488
Qy       1517 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 1576
```

```
Db        489 CACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGT 548

Qy       1577 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 1636
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 GATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTA 608

Qy       1637 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 1696
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 CAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTG 668

Qy       1697 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTC 1756
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTC 728

Qy       1757 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 1816
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 AACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAAC 788

Qy       1817 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 1876
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAA 848

Qy       1877 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 1936
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTT 908

Qy       1937 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 1996
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 CTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACAT 968

Qy       1997 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 2056
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 GAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGAC 1028

Qy       2057 CCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 2116
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 CCTACAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACA 1088

Qy       2117 GATATTCA 2124
             ||||||| |
Db       1089 GATATTTA 1096


RESULT 26
US-10-199-937-138
; Sequence 138, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 138
;   LENGTH: 5482
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (474)..(4016)
US-10-199-937-138

  Query Match             26.6%;  Score 1008.8;  DB 14;  Length 5482;
  Best Local Similarity   58.2%;  Pred. No. 9.2e-259;

  Matches 1955;  Conservative    0;  Mismatches 1292;  Indels  111;  Gaps    6;

Qy        404 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCA 463
              |||| ||| ||    || || |||||||   |||  |  || ||     | || ||
Db        669 GGTTATGGACGCCGGGAAGTGGTTGAATTCCTGCTGAACAGCGGCGCCTCCATACAGGCG 728

Qy        464 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 523
              |||| || || || ||     || || || ||    ||||| ||||| || || || |||
Db        729 TGTGACGAGGGTGGGCTGCACCCGCTGCACAACTGTTGCTCCTTTGGCCACGCCGAGGTA 788

Qy        524 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 583
              ||   |  |  ||     |||||    |||  |  || || ||||| || ||
Db        789 GTTCGATTGTTGCTGAAGGCAGGTGCCAGTCCAAACACCACCGACAACTGGAACTACACG 848

Qy        584 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 643
              ||  | || || || || |  || || ||| | ||||| ||| | |  ||||| ||||||
Db        849 CCATTGCACGAGGCGGCCAGCAAGGGCAAGGTGGATGTGTGCCTGGCTCTGTTGCAGCAT 908

Qy        644 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 703
              || ||  | |  || ||||| ||  | ||    | |||| || |||| | || ||
Db        909 GGCGCAAACCATACGATCCGCAACTCGGAGCAGAAGACACCACTGGAGCTGGCGGACGAG 968

Qy        704 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 763
              |  |    || | || || |||||   || ||||| ||  | |||   ||  |
Db        969 GCGACGCGTCCCGTATTGACCGGCGAATATCGAAAGGATGAGCTGCTTGAAGCCGCACGC 1028

Qy        764 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 823
              ||    || |||| |||||||| ||| | |||||||||| || ||
Db       1029 TCGGGGGCCGAGGATCGCCTGCTGGCCCTACTCACGCCACTCAATGTCAACTGTCATGCC 1088

Qy        824 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883
              || |||||  ||  |||||  ||  | ||| |||||| || ||  ||  |  ||
Db       1089 AGCGATGGACGACGCTCAACGCCGCTCCATCTGGCAGCGGGCTACAATCGGATCGGCATC 1148

Qy        884 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 943
              ||  |  |  ||| | |||||  ||||| |||||||| || ||  |  ||||||
Db       1149 GTGGAAATTCTGCTGGCCAACGGAGCGGATGTACATGCTAAGGACAAGGGCGGTCTGGTG 1208

Qy        944 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1003
              ||  | ||||||||||| || || || || | || || ||  |  ||  || ||
Db       1209 CCGCTGCACAATGCCTGCTCCTACGGACACTTCGATGTGACCAAGCTGCTTATCCAGGCG 1268

Qy       1004 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1063
              || |||  ||| || || |  ||  ||||  || || || || || ||||| || ||
Db       1269 GGCGCCAATGTCAACGCCAACGATCTGTGGGCCTTTACGCCGCTCCACGAGGCCGCCTCC 1328

Qy       1064 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1123
              || |  | || || || ||  || |  | |||  ||| || || || || || || ||
Db       1329 AAAAGTCGCGTCGAGGTCTGCAGCCTGCTGCTCAGTCGTGGAGCGGATCCCACCCTCCTA 1388

Qy       1124 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1183
              || || |||| ||  || || ||  |||  |||||  || | | ||| |  |  |
Db       1389 AACTGCCACAGCAAGTCGGCCATCGATGCGGCGCCCACCAGGGAGCTGAGAGAGCGGATT 1448

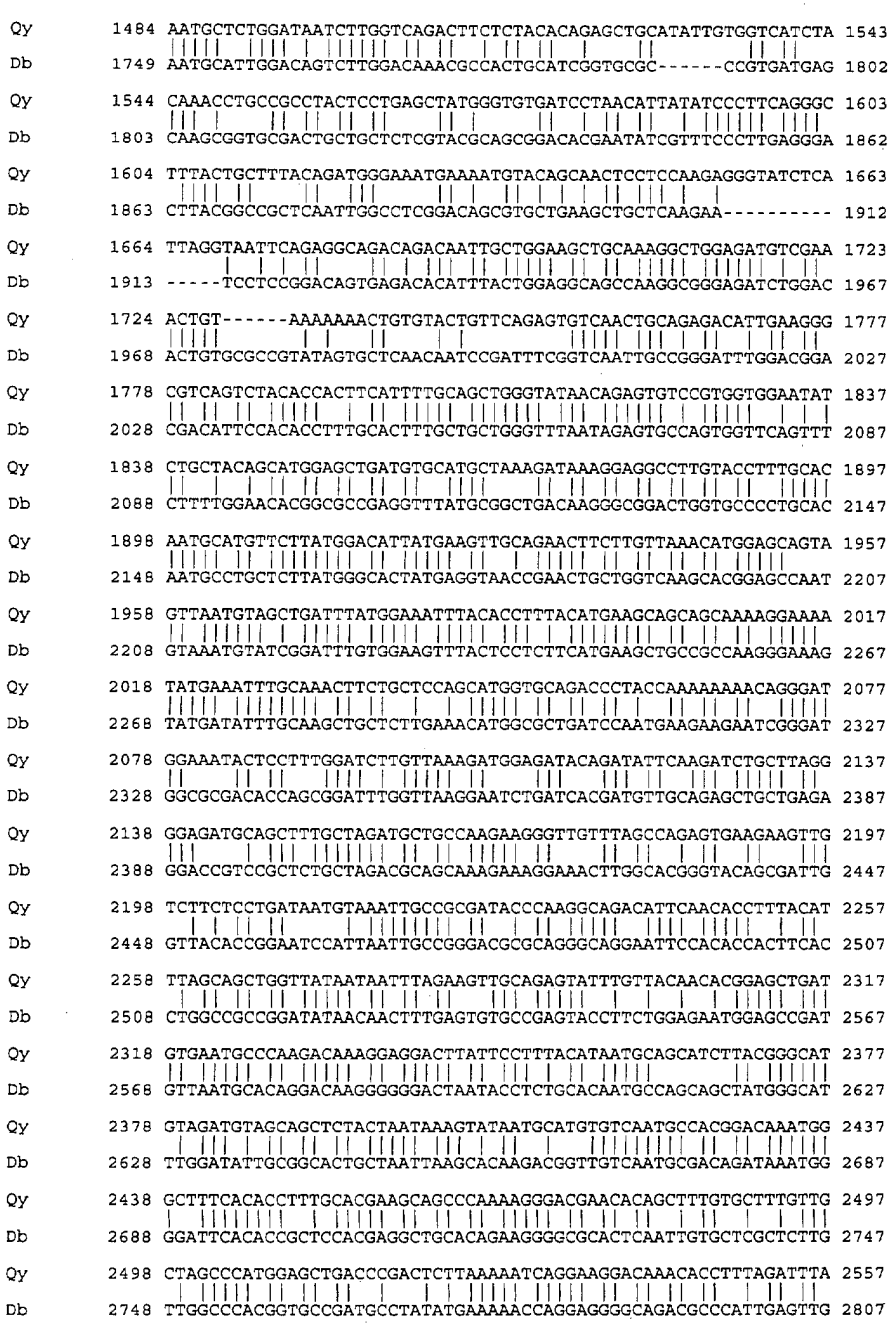
Qy       1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1243
              || | |||||  || || ||||  |||||  | ||  ||| |  |  ||||||  | ||
Db       1449 GCCTTTGAATACAAGGGTCACTGCCTGCTGGACGCCTGTCGAAAGTGTGATGTGTCCCGT 1508

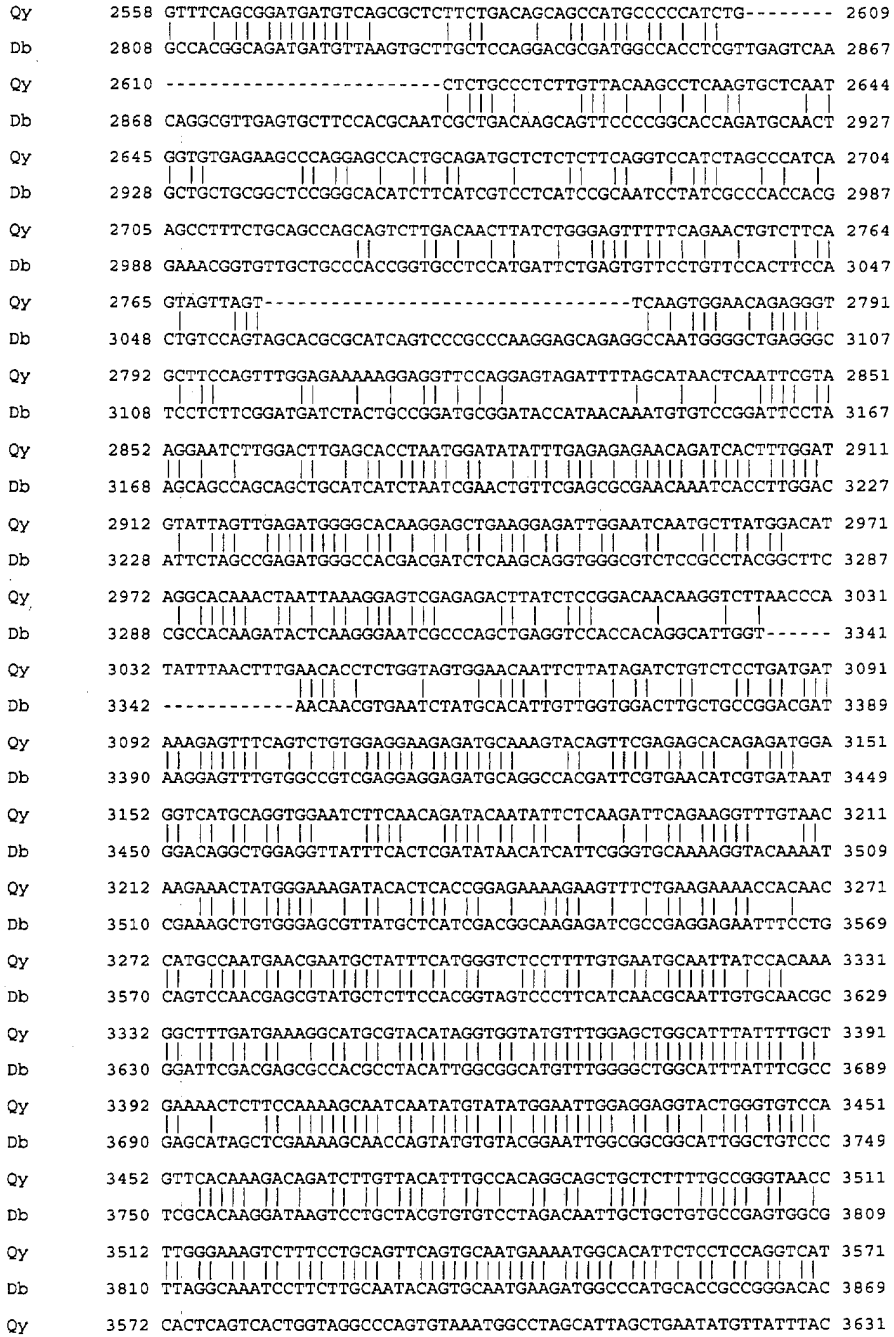
Qy       1244 ATCAAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1303
                ||| || |  | |  || || || || ||| ||||||  | |||  || ||
Db       1509 GCCAAGAAGCTGGTATGCGCAGAGATTGTTAACTTCGTGCATCCATATACAGGAGACACT 1568

Qy       1304 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1363
              |  | ||   || | ||  ||| ||  ||  | |||||  |   ||||||  ||
Db       1569 CCGCTCCACCTGGCCGTTGTCAGTCCGGATGGGAAGCGCAAGCAGCTGATGGAACTGCTG 1628

Qy       1364 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1423
              |||||  |||  |  | || || || |  || |  ||| | || || |||||  ||
Db       1629 ACCAGAAAGGGATCCTTGCTGAACGAGAAAAACAAGGCTTTCCTCACGCCCCTGCATTTG 1688

Qy       1424 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1483
              ||  | |||   |||  | ||||  | || ||  ||  | || ||  | || ||||||
Db       1689 GCTGCCGAGCTGCTTCACTACGATGCCATGGAGGTGCTGCTAAAGCAGGGCGCCAAGGTT 1748
```





```
Db       3870 CACTCGGTGGTGGGCAGACCCTCGGCGGGTGGCTTGCATTTCGCCGAATACGTTGTCTAT 3929

Qy       3632 AGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3689

Db       3930 CGGGGCGAACAGTCTTATCCGGAGTACTTGATAACCTACCAAATCGTCAAGCCCGATG 3987


RESULT 27
US-09-964-899-40
; Sequence 40, Application US/09964899
; Patent No. US20020174446A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Dalia et al.
;  TITLE OF INVENTION: Identification of Genes Involved in
;  TITLE OF INVENTION:  Alzheimer's Disease Using Drosophila Melanogaster
;  FILE REFERENCE: 4-31612 A
;  CURRENT APPLICATION NUMBER: US/09/964,899
;  CURRENT FILING DATE:  2001-09-27
;  PRIOR APPLICATION NUMBER: 60/236,893
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/298,309
;  PRIOR FILING DATE: 2001-06-14
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 2409
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-964-899-40

  Query Match             26.2%;  Score 994;  DB 9;  Length 2409;
  Best Local Similarity   65.4%;  Pred. No. 4.9e-255;
  Matches 2174;  Conservative    0;  Mismatches   20;  Indels 1128;  Gaps    4;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462

Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522

Db        258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582

Db        318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642

Db        378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702

Db        438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATAT----------------AAGAAAGATGA 744

Db        498 ATCTGCCAAAGCAGTGCTTACTGGTAAGTCTGTATACTCTGGTTATTCCAGGAAGCCTGT 557

Qy        745 ACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATT 804

Db        558 AAAGAACAACCTTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATT 617

Qy        805 AAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGG 864

Db        618 AAATGTCAACTGCCACGCAAGTGATGGCAGA---------------------------- 648

Qy        865 ATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAA 924

Db        649 ------------------------------------------------------------ 648

Qy        925 AGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAAC 984

Db        649 ------------------------------------------------------------ 648

Qy        985 TGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCC 1044

Db        649 ------------AAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCC 695

Qy       1045 TCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGG 1104

Db        696 TCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGG 755

Qy       1105 TGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACC 1164

Db        756 TGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACC 815

Qy       1165 ACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACG 1224

Db        816 ACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACG 875

Qy       1225 AGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCA 1284

Db        876 AGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCA 935

Qy       1285 TCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAA 1344

Db        936 TCCTCAAACACATGAAACAGCATTG---------------------------------- 960

Qy       1345 GCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATT 1404

Db        961 ------------------------------------------------------------ 960

Qy       1405 CTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGT 1464

Db        961 ------------------------------------------------------------ 960

Qy       1465 GAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGC 1524

Db        961 ------------------------------------------------------------ 960

Qy       1525 TGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAA 1584

Db        961 ------------------------------------------------------------ 960

Qy       1585 CATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACT 1644

Db        961 ------------------------------------------------------------ 960

Qy       1645 CCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGC 1704

Db        961 ------------------------------------------------------------ 960

Qy       1705 AAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAG 1764

Db        961 ----------------------------AAACTGTGTACTGTTCAGAGTGTCAACTGCAG 992

Qy       1765 AGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTC 1824

Db        993 AGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTC 1052

Qy       1825 CGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCT 1884

Db       1053 CGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCT 1112

Qy       1885 TGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAA 1944

Db       1113 TGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAA 1172

Qy       1945 ACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGC 2004

Db       1173 ACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGC 1232

Qy       2005 AGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAA 2064

Db       1233 AGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAA 1292
```

```
Qy       2065 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 2124
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1293 AAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCA 1352

Qy       2125 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 2184
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1353 AGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAG 1412

Qy       2185 AGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTC 2244
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1413 AGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTC 1472

Qy       2245 AACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACA 2304
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1473 AACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACA 1532

Qy       2305 ACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGC 2364
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1533 ACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGC 1592

Qy       2365 ATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGC 2424
             ||||||||
Db       1593 ATCTTACGG---------------------------------------------------- 1601

Qy       2425 CACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCT 2484
Db       1602 ---------------------------------------------------------- 1601

Qy       2485 TTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAAC 2544
Db       1602 ---------------------------------------------------------- 1601

Qy       2545 ACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCC 2604
Db       1602 ---------------------------------------------------------- 1601

Qy       2605 ATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGC 2664
Db       1602 ---------------------------------------------------------- 1601

Qy       2665 CACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAG 2724
Db       1602 ---------------------------------------------------------- 1601

Qy       2725 TCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAAC 2784
Db       1602 ---------------------------------------------------------- 1601

Qy       2785 AGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCA 2844
Db       1602 ---------------------------------------------------------- 1601

Qy       2845 ATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCAC 2904
                                                                  ||||||
Db       1602 ----------------------------------------------------GATCAC 1607

Qy       2905 TTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 2964
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1608 TTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTA 1667

Qy       2965 TGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCT 3024
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1668 TGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCT 1727

Qy       3025 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 3084
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1728 TAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCC 1787

Qy       3085 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 3144
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1788 TGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAG 1847

Qy       3145 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 3204
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1848 AGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGT 1907

Qy       3205 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 3264
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1908 TTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAA 1967

Qy       3265 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 3324
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1968 CCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTAT 2027

Qy       3325 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 3384
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2028 CCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTA 2087

Qy       3385 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 3444
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2088 TTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGG 2147

Qy       3445 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 3504
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2148 GTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCG 2207

Qy       3505 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 3564
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2208 GGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCC 2267

Qy       3565 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 3624
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2268 AGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGT 2327

Qy       3625 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 3684
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2328 TATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCC 2387

Qy       3685 TGAAGGTATGGTCGATGGATAA 3706
             ||||||||||||||||||||||
Db       2388 TGAAGGTATGGTCGATGGATAA 2409
```

RESULT 28
US-10-252-157-341
; Sequence 341, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 341
; LENGTH: 797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 405041.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 257
; OTHER INFORMATION: a, t, c, g, or other
US-10-252-157-341

Query Match 20.9%; Score 794.4; DB 14; Length 797;
Best Local Similarity 99.7%; Pred. No. 7.2e-202;
Matches 795; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1344 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1403
Db          1 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 60

Qy       1404 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1463
Db         61 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 120

Qy       1464 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1523
Db        121 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 180

Qy       1524 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1583
Db        181 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 240

Qy       1584 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 1643
Db        241 ACATTATATCCCTTCANGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAAC 300

Qy       1644 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 1703
Db        301 TCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTG 360

Qy       1704 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 1763
Db        361 CAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCA 420

Qy       1764 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 1823
Db        421 GAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGT 480

Qy       1824 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 1883
Db        481 CCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCC 540

Qy       1884 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 1943
Db        541 TTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTA 600

Qy       1944 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 2003
Db        601 AACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAG 660

Qy       2004 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCA 2063
Db        661 CAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAA 720

Qy       2064 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 2123
Db        721 AAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTC 780

Qy       2124 AAGATCTGCTTAGGGGA 2140
Db        781 AAGATCTGCTTAGGGGA 797
```

RESULT 29
US-10-199-937-45
; Sequence 45, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence not
; OTHER INFORMATION: specified as protein-coding is vector sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(480)
US-10-199-937-45

Query Match 15.1%; Score 574; DB 14; Length 582;
Best Local Similarity 100.0%; Pred. No. 6.9e-143;
Matches 574; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3224 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3283
Db          1 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 60

Qy       3284 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3343
Db         61 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 120

Qy       3344 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3403
Db        121 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 180

Qy       3404 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3463
Db        181 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 240

Qy       3464 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3523
Db        241 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 300

Qy       3524 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3583
Db        301 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 360

Qy       3584 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643
Db        361 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 420

Qy       3644 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
Db        421 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 480

Qy       3704 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3763
Db        481 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 540

Qy       3764 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
Db        541 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAA 574
```

RESULT 30
US-09-972-115A-10/c
; Sequence 10, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05

```
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 520
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-10

  Query Match             13.7%;  Score 518.4;  DB 10;  Length 520;
  Best Local Similarity   99.8%;  Pred. No. 5.1e-128;
  Matches  519;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        658 CATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGT 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 CATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGT 461

Qy        718 GCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGA 777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGA 401

Qy        778 AAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAA 837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 AAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAA 341

Qy        838 GTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACT 897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACT 281

Qy        898 GCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGC 957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGC 221

Qy        958 CTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAA 1017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 CTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAA 161

Qy       1018 TGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGA 1077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 TGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGA 101

Qy       1078 AGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAA 1137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAA 41

Qy       1138 AAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAA 1177
              ||||||||||||||||||||||||||||||||||||||||| ||
Db          40 AAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAAAA 1


RESULT 31
US-09-833-381-841
; Sequence 841, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 841
;   LENGTH: 1069
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1069)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-841

  Query Match             12.3%;  Score 467.6;  DB 9;  Length 1069;
  Best Local Similarity   98.8%;  Pred. No. 3.4e-114;
  Matches  492;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
              | | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CCCTTCCTAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 626

Qy       1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 686

Qy       1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 746

Qy       1823 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 806

Qy       1883 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 866

Qy       1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 926

Qy       2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACA 986

Qy       2063 AAAAAAAACA-GGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATAT 2121
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        987 AAAAAAAACAGGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGAT-CAGATAT 1045

Qy       2122 TCAAGATCTGCTTAGGGG 2139
              ||||||||||||||||||
Db       1046 TCAAGATCTGCTTAGGGG 1063


RESULT 32
US-09-833-381-1153
; Sequence 1153, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1153
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1153

  Query Match             12.1%;  Score 461;  DB 9;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-112;
  Matches  461;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        553 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 612
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 64
```

```
Qy        613 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 124

Qy        673 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 184

Qy        733 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 244

Qy        793 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 304

Qy        853 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 364

Qy        913 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 424

Qy        973 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 1013
              |||||||||||||||||||||||||||||||||||||||||
Db        425 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 465
```

RESULT 33
US-10-106-698-2223
; Sequence 2223, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2223
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (136)..(136)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2223

Query Match 12.1%; Score 459; DB 14; Length 460;
Best Local Similarity 99.8%; Pred. No. 3.9e-112;
Matches 459; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2753 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 2812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 60

Qy       2813 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 2872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 120

Qy       2873 CTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 2932
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 180

Qy       2933 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 2992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 240

Qy       2993 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 3052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 300

Qy       3053 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 3112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 360

Qy       3113 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 3172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 420

Qy       3173 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 3212
              |||||||||||||||||||||||||||||||||||||||
Db        421 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 460
```

RESULT 34
US-10-199-937-5
; Sequence 5, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (136)
; OTHER INFORMATION: n= a, c, g, or t
US-10-199-937-5

Query Match 12.1%; Score 459; DB 14; Length 460;
Best Local Similarity 99.8%; Pred. No. 3.9e-112;
Matches 459; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2753 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 2812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG 60

Qy       2813 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 2872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCAC 120

Qy       2873 CTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 2932
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTAATGGATATATTTNAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCAC 180

Qy       2933 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 2992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGA 240

Qy       2993 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 3052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCT 300
```

```
Qy       3053 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 3112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAG 360

Qy       3113 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 3172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTC 420

Qy       3173 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 3212
              |||||||||||||||||||||||||||||||||||||||
Db        421 AACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACA 460


RESULT 35
US-10-181-447A-11
; Sequence 11, Application US/10181447A
; Publication No. US20030180738A1
; GENERAL INFORMATION:
;  APPLICANT: The No. US20030180738A1tingham Trent University
;  TITLE OF INVENTION: Cancer Associated Genes and their Products
;  FILE REFERENCE: No. US20030180738A1tingham Trent Uni
;  CURRENT APPLICATION NUMBER: US/10/181,447A
;  CURRENT FILING DATE:  2002-07-18
;  PRIOR APPLICATION NUMBER: PCT/GB/01/000188
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: GB0000993.6
;  PRIOR FILING DATE: 2000-01-18
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 402
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-181-447A-11

  Query Match             10.5%;  Score 396.8;  DB 14;  Length 402;
  Best Local Similarity   99.5%;  Pred. No. 1.7e-95;
  Matches  398;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        972 ATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGT 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGT 60

Qy       1032 GGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTC 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTC 120

Qy       1092 TCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACT 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACT 180

Qy       1152 TGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGC 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGC 240

Qy       1212 TGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGG 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGG 300

Qy       1272 TGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCAT 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCAT 360

Qy       1332 ATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAA 1371
              |||||||||||||||||||||||||||||||||||  |||
Db        361 ATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAGAAAA 400


RESULT 36
US-09-972-115A-9
; Sequence 9, Application US/09972115A
; Publication No. US20030032769A1
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-9

  Query Match             10.0%;  Score 380.4;  DB 10;  Length 384;
  Best Local Similarity   99.7%;  Pred. No. 4e-91;
  Matches  381;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1682 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 62

Qy       1742 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 122

Qy       1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 182

Qy       1862 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 242

Qy       1922 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 302

Qy       1982 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 362

Qy       2042 CAGCATGGTGCAGACCCTACCA 2063
              |||||||||||||||||||| |
Db        363 CAGCATGGTGCAGACCCTACAA 384


RESULT 37
US-09-918-995-17041
; Sequence 17041, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17041
;   LENGTH: 441
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(441)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17041

  Query Match             9.7%;  Score 368;  DB 10;  Length 441;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-88;
  Matches  368;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3427 AATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAG 3486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAG 60

Qy       3487 GCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAA 3546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAA 120

Qy       3547 AATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCT 3606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCT 180

Qy       3607 AGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAC 3666
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTAC 240

Qy       3667 TTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTA 3726
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTA 300

Qy       3727 ATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 360

Qy       3787 AAAAAAAA 3794
             ||||||||
Db        361 AAAAAAAA 368


RESULT 38
US-10-199-937-41
; Sequence 41, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 1691
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(357)
US-10-199-937-41

  Query Match             9.5%;  Score 359.6;  DB 14;  Length 1691;
  Best Local Similarity   90.8%;  Pred. No. 4.2e-85;
  Matches  383;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       3287 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 60

Qy       3347 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3406
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 120

Qy       3407 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3466
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 180

Qy       3467 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3526
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 240

Qy       3527 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3586
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 300

Qy       3587 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGTA 360

Qy       3647 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAA 3706
             ||||| | | | | | | ||| | | | | | | | | | | | | | | | | | | | ||
Db        361 ATGTAGTTTTATTTGTTCATCTTCAAAAATGCTAGGGAGGCATACTTTAACTTTTTATTA 420

Qy       3707 AT 3708
             ||
Db        421 AT 422


RESULT 39
US-10-199-937-43
; Sequence 43, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
;   LENGTH: 1692
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Sequence not
;   OTHER INFORMATION: specified as protein-coding is vector sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(357)
US-10-199-937-43

  Query Match             9.5%;  Score 359.6;  DB 14;  Length 1692;
  Best Local Similarity   90.8%;  Pred. No. 4.2e-85;
  Matches  383;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       3287 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 3346
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGG 60
```

```
Qy       3347 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 3406
Db         61 CATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAA 120

Qy       3407 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 3466
Db         121 AGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGA 180

Qy       3467 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 3526
Db         181 TCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTC 240

Qy       3527 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 3586
Db         241 CTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGT 300

Qy       3587 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCT 3646
Db         301 AGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGTA 360

Qy       3647 TATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAA 3706
Db         361 ATGTAGTTTTATTTGTTCATCTTCAAAAATGCTAGGGAGGCATACTTTAACTTTTTATTA 420

Qy       3707 AT 3708
Db         421 AT 422
```

RESULT 40
US-10-199-937-49
; Sequence 49, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-937-49

Query Match 9.4%; Score 356; DB 14; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.3e-84;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       3008 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 3067
Db          1 TCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATT 60

Qy       3068 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 3127
Db            61 CTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGT 120

Qy       3128 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 3187
Db         121 ACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATT 180

Qy       3188 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 3247
Db         181 CTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAA 240

Qy       3248 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 3307
Db         241 GAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCT 300

Qy       3308 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGG 3363
Db         301 TTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGG 356
```

RESULT 41
US-10-199-937-7
; Sequence 7, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
; APPLICANT: Christenson, Erik
; APPLICANT: DeMaggio, Anthony J.
; APPLICANT: Goldman, Phyllis S.
; APPLICANT: McElligott, David L.
; TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
; FILE REFERENCE: 27866/36559
; CURRENT APPLICATION NUMBER: US/10/199,937
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/606,035
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/141,582
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (203)
; OTHER INFORMATION: n= a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (388)
; OTHER INFORMATION: n= a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (441)
; OTHER INFORMATION: n= a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (456)
; OTHER INFORMATION: n= a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (538)
; OTHER INFORMATION: n= a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (543)
; OTHER INFORMATION: n= a, c, g, or t
US-10-199-937-7

Query Match 9.3%; Score 354.6; DB 14; Length 564;
Best Local Similarity 93.4%; Pred. No. 4.3e-84;
Matches 369; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy       3288 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 3347
Db          1 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 60

Qy       3348 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 3407
Db            61 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 120

Qy       3408 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 3467
```

```
Db        121 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 180

Qy       3468 CTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 3527
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||
Db        181 CTTGTTACATTTGCCACAGGAGNCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 240

Qy       3528 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 3587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 300

Qy       3588 GGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTT 3647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGTAA 360

Qy       3648 ATCCTGAGTATTTAATTACTTACCAGATTATGAGG 3682
                  |||||  |  |  |  |  |  |   |||
Db        361 TGTAGTTTTATTTGTTCATCTTCAAAANTGCTAGG 395


RESULT 42
US-10-199-937-14
; Sequence 14, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 392
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (331)
;   OTHER INFORMATION: n= a, c, g, or t
US-10-199-937-14

  Query Match             9.1%;  Score 346;  DB 14;  Length 392;
  Best Local Similarity   99.7%;  Pred. No. 6.8e-82;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3288 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 3347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGC 60

Qy       3348 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 3407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAA 120

Qy       3408 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 3467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGAT 180

Qy       3468 CTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 3527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCC 240

Qy       3528 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 3587
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTA 300

Qy       3588 GGCCCAGTGTAAATGGCCTAGCATTAGCTG-AATATGTTATTTACAGAGGAGAACAGG 3644
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        301 GGCCCAGTGTAAATGGCCTAGCATTAGCTGNAATATGTTATTTACAGAGGAGAACAGG 358


RESULT 43
US-10-199-937-12/c
; Sequence 12, Application US/10199937
; Publication No. US20030190739A1
; GENERAL INFORMATION:
;  APPLICANT: Christenson, Erik
;  APPLICANT:  DeMaggio, Anthony J.
;  APPLICANT:  Goldman, Phyllis S.
;  APPLICANT:  McElligott, David L.
;  TITLE OF INVENTION: TANKYRASE2 MATERIALS AND METHODS
;  FILE REFERENCE: 27866/36559
;  CURRENT APPLICATION NUMBER: US/10/199,937
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/606,035
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/141,582
;  PRIOR FILING DATE: 1999-06-29
;  NUMBER OF SEQ ID NOS: 178
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-199-937-12

  Query Match             9.0%;  Score 340.4;  DB 14;  Length 353;
  Best Local Similarity   99.4%;  Pred. No. 2e-80;
  Matches  352;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       3433 AGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCT 3492
              || ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        353 AGAAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTT-CATTTGCCACAGGCAGCT 295

Qy       3493 GCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 3552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 GCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGC 235

Qy       3553 ACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATT 3612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        234 ACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATT 175

Qy       3613 AGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCA 3672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        174 AGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCA 115

Qy       3673 GATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCA 3732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        114 GATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCA 55

Qy       3733 CTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 3786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          54 CTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTG 1


RESULT 44
US-10-040-739-1179
; Sequence 1179, Application US/10040739
; Publication No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
```

```
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;       COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 1179:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 353 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1179:
US-10-040-739-1179

  Query Match             8.5%;  Score 323;  DB 13;  Length 353;
  Best Local Similarity   99.7%;  Pred. No. 9.2e-76;
  Matches  323;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3233 ACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTA 3292
             ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db         21 ACTCACCGGAGAAAAGAAGTTTCNGAAGAAAACCACAACCATGCCAATGAACGAATGCTA 80

Qy       3293 TTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCG 3352
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 TTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCG 140

Qy       3353 TACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAAT 3412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 TACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAAT 200

Qy       3413 CAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGT 3472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 CAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGT 260

Qy       3473 TACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAG 3532
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 TACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAG 320

Qy       3533 TTCAGTGCAATGAAAATGGCACAT 3556
             |||||||||||||||||||||||
Db        321 TTCAGTGCAATGAAAATGGCACAT 344


RESULT 45
US-10-085-783A-34884
; Sequence 34884, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34884
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-34884

  Query Match             7.5%;  Score 285.6;  DB 12;  Length 471;
  Best Local Similarity   76.7%;  Pred. No. 1.2e-65;
  Matches  362;  Conservative    0;  Mismatches  109;  Indels    1;  Gaps    1;

Qy       3161 GGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTA 3220
             || || |||||||||||||||||| | |    ||||| || |||   ||||||||  |
Db          1 GGCGGCATCTTCAACAGATACAATGTCATTCGAATTCA-AAAGTTGTCAACAAGAAGTTG 59

Qy       3221 TGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAAT 3280
              ||||  | | |  |||||     || ||||| ||||| || || |||||||||  ||||
Db         60 AGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAAT 119

Qy       3281 GAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGAT 3340
              || || ||||| |||||||| ||||||||  | |||||  ||||| || |||||  ||||||
Db        120 GAGCGCATGCTGTTTCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGAT 179

Qy       3341 GAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCT 3400
              ||  | |||||  |||||||| || |||||||| || || ||||||||||||||||||||
Db        180 GAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCC 239

Qy       3401 TCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAA 3460
              || |||||||| |||||||| ||||||||||||||||||| || || || ||    |||||
Db        240 TCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAG 299

Qy       3461 GACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAG 3520
              ||||| || || || || || ||||| || |||||| ||  | || ||| | || ||
Db        300 GACAGGTCATGCTATATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAA 359

Qy       3521 TCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTC 3580
              || || |||||||| ||  | |||||||||| ||  | |||||||| || |||||||||
Db        360 TCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTC 419

Qy       3581 ACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACA 3632
              | |||||| || || || ||||| || ||||  |||||||||| || ||||
Db        420 ATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAATATGTCATCTACA 471


Search completed: March 30, 2004, 17:11:12
Job time : 1334.29 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2004, 19:31:32 ; Search time 247.88 Seconds
(without alignments)
8500.686 Million cell updates/sec

Title: US-09-843-159B-1
Perfect score: 3797
Sequence: 1 ctttgaagacactggatttc.....cctttgctgaaaaaaaaaaa 3797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3797	100.0	3797	4	US-09-696-668-1	Sequence 1, Appli
	2	3393.4	89.4	3816	4	US-09-696-668-2	Sequence 2, Appli
	3	3388.8	89.2	4275	4	US-09-972-115A-5	Sequence 5, Appli
	4	3386	89.2	3394	4	US-09-427-154-1	Sequence 1, Appli
	5	3382.8	89.1	4512	4	US-09-350-982C-3	Sequence 3, Appli
	6	3291.8	86.7	3498	4	US-09-350-982C-4	Sequence 4, Appli
	7	2969.6	78.2	4493	4	US-09-972-115A-1	Sequence 1, Appli
	8	2968.2	78.2	4297	4	US-09-972-115A-3	Sequence 3, Appli
	9	1813.4	47.8	4134	3	US-09-196-387-1	Sequence 1, Appli
	10	1813.4	47.8	4134	4	US-09-841-835-1	Sequence 1, Appli
	11	1813.4	47.8	4134	4	US-09-972-115A-7	Sequence 7, Appli
	12	1449.6	38.2	4491	3	US-09-196-387-7	Sequence 7, Appli
	13	1449.6	38.2	4491	4	US-09-841-835-7	Sequence 7, Appli
	14	1297.8	34.2	4657	3	US-09-196-387-9	Sequence 9, Appli
	15	1297.8	34.2	4657	4	US-09-841-835-9	Sequence 9, Appli
c	16	518.4	13.7	520	4	US-09-972-115A-10	Sequence 10, Appl
	17	467.6	12.3	1069	4	US-09-833-381-841	Sequence 841, App
	18	461	12.1	465	4	US-09-833-381-1153	Sequence 1153, Ap
	19	380.4	10.0	384	4	US-09-972-115A-9	Sequence 9, Appli
	20	121.2	3.2	523	3	US-09-196-387-12	Sequence 12, Appl
	21	121.2	3.2	523	4	US-09-841-835-12	Sequence 12, Appl
	22	97.4	2.6	14770	4	US-09-220-132-30	Sequence 30, Appl
	23	93	2.4	5235	1	US-09-031-485-35	Sequence 35, Appl
c	24	93	2.4	5235	1	US-09-031-485-36	Sequence 36, Appl
	25	93	2.4	5235	1	US-08-847-429A-35	Sequence 35, Appl
c	26	93	2.4	5235	1	US-08-847-429A-36	Sequence 36, Appl
	27	93	2.4	5235	3	US-09-065-474-35	Sequence 35, Appl
c	28	93	2.4	5235	3	US-09-065-474-36	Sequence 36, Appl
	29	93	2.4	5235	4	US-09-557-034-35	Sequence 35, Appl
c	30	93	2.4	5235	4	US-09-557-034-36	Sequence 36, Appl
	31	93	2.4	5503	1	US-09-031-485-32	Sequence 32, Appl
c	32	93	2.4	5503	1	US-09-031-485-34	Sequence 34, Appl
	33	93	2.4	5503	1	US-08-847-429A-32	Sequence 32, Appl
c	34	93	2.4	5503	1	US-08-847-429A-34	Sequence 34, Appl
	35	93	2.4	5503	3	US-09-065-474-32	Sequence 32, Appl
c	36	93	2.4	5503	3	US-09-065-474-34	Sequence 34, Appl
	37	93	2.4	5503	4	US-09-557-034-32	Sequence 32, Appl
c	38	93	2.4	5503	4	US-09-557-034-34	Sequence 34, Appl
	39	83.4	2.2	1056	3	US-09-065-474-138	Sequence 138, App
c	40	83.4	2.2	1056	3	US-09-065-474-140	Sequence 140, App
	41	83.4	2.2	1056	4	US-09-557-034-138	Sequence 138, App
c	42	83.4	2.2	1056	4	US-09-557-034-140	Sequence 140, App
c	43	81.6	2.1	466	4	US-09-833-381-842	Sequence 842, App
	44	78.8	2.1	3454	3	US-09-082-059-1	Sequence 1, Appli
	45	74.2	2.0	2505	3	US-09-291-839-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-696-668-1
; Sequence 1, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Chan, Eva
; APPLICANT: Xu, Xiang
; APPLICANT: Huang, Betty
; APPLICANT: Ossovskaya, Valeria
; TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS (
; FILE REFERENCE: A-68292-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/696,668
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/427,154
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3797
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-696-668-1

Query Match 100.0%; Score 3797; DB 4; Length 3797;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTTTGAAGACACTGGATTTCATACTTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTTTGAAGACACTGGATTTCATACTTTTGCCTGGGGTTATCTCTCTGTGTCTCACTACAT 60

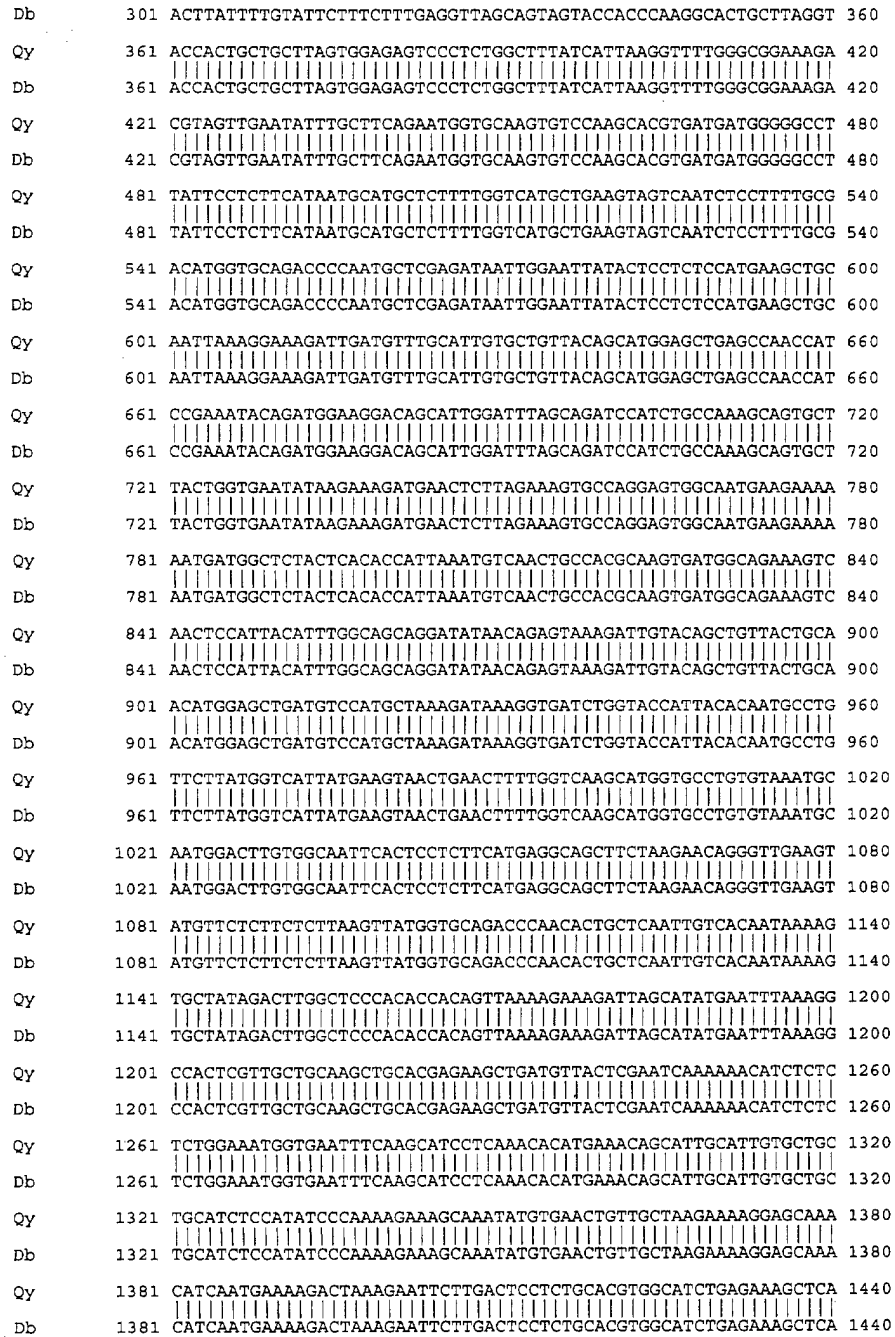
Qy         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACAAATATTAGCTGTGAGCAGATCTTTTTTTGTTGCTTCTTGTAGTCCCCCAGTTTAG 120

Qy        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CAGAAACATTCTGTGAGATAGATGTGGGAAAGGAATTCTAGCAAGAGTTTTGTCACTGTA 180

Qy        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCATAAGGTTGTGATTTACATATTTAAGTTTTATACTTTGAACATCTGAAAATGTATACA 240

Qy        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TACTAAATATGCAGAACTCTATTGTAGAGTGAGAAACATTTGAACTTTGAGCTTTCAGTC 300

Qy        301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 301 ACTTATTTTGTATTCTTTCTTTGAGGTTAGCAGTAGTACCACCCAAGGCACTGCTTAGGT 360
Qy 361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420
Db 361 ACCACTGCTGCTTAGTGGAGAGTCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGA 420
Qy 421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480
Db 421 CGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCT 480
Qy 481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540
Db 481 TATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCG 540
Qy 541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600
Db 541 ACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGC 600
Qy 601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660
Db 601 AATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCAT 660
Qy 661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720
Db 661 CCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCT 720
Qy 721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780
Db 721 TACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAA 780
Qy 781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840
Db 781 AATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTC 840
Qy 841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900
Db 841 AACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCA 900
Qy 901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960
Db 901 ACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTG 960
Qy 961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020
Db 961 TTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGC 1020
Qy 1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080
Db 1021 AATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGT 1080
Qy 1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140
Db 1081 ATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAG 1140
Qy 1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200
Db 1141 TGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGG 1200
Qy 1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260
Db 1201 CCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTC 1260
Qy 1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320
Db 1261 TCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTGCTGC 1320
Qy 1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380
Db 1321 TGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAA 1380
Qy 1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440
Db 1381 CATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCA 1440
Qy 1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500
Db 1441 TAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCT 1500
Qy 1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560
Db 1501 TGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACT 1560
Qy 1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620
Db 1561 CCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCTTTACAGAT 1620
Qy 1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
Db 1621 GGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGC 1680
Qy 1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740
Db 1681 AGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTG 1740
Qy 1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800
Db 1741 TACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTT 1800
Qy 1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860
Db 1801 TGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGT 1860
Qy 1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
Db 1861 GCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTA 1920
Qy 1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
Db 1921 TGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAA 1980
Qy 1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
Db 1981 ATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCT 2040
Qy 2041 CCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Db 2041 CCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGT 2100
Qy 2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Db 2101 TAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGC 2160
Qy 2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Db 2161 TGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTG 2220
Qy 2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Db 2221 CCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTT 2280
Qy 2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Db 2281 AGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGG 2340
Qy 2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Db 2341 ACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAAT 2400
Qy 2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Db 2401 AAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGC 2460
Qy 2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520
Db 2461 AGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGAC 2520

```
Qy       2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580
Db       2521 TCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGC 2580

Qy       2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640
Db       2581 TCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCT 2640

Qy       2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Db       2641 CAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700

Qy       2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Db       2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760

Qy       2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820
Db       2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCC 2820

Qy       2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880
Db       2821 AGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGA 2880

Qy       2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940
Db       2881 TATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCT 2940

Qy       2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000
Db       2941 GAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAG 3000

Qy       3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060
Db       3001 ACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGG 3060

Qy       3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120
Db       3061 AACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGAT 3120

Qy       3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180
Db       3121 GCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATA 3180

Qy       3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240
Db       3181 CAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCG 3240

Qy       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300
Db       3241 GAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGG 3300

Qy       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360
Db       3301 GTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGG 3360

Qy       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420
Db       3361 TGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGT 3420

Qy       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480
Db       3421 ATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTG 3480

Qy       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540
Db       3481 CCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGC 3540

Qy       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600
Db       3541 AATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAA 3600

Qy       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660
Db       3601 TGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTT 3660

Qy       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720
Db       3661 AATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAG 3720

Qy       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780
Db       3721 AAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCT 3780

Qy       3781 TTGCTGAAAAAAAAAAA 3797
Db       3781 TTGCTGAAAAAAAAAAA 3797
```

RESULT 2
US-09-696-668-2
; Sequence 2, Application US/09696668
; Patent No. 6617102
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Chan, Eva
; APPLICANT: Xu, Xiang
; APPLICANT: Huang, Betty
; APPLICANT: Ossovskaya, Valeria
; TITLE OF INVENTION: TANKYRASE H, COMPOSITIONS INVOLVED IN THE CELL CYCLE AND METHODS (
; FILE REFERENCE: A-68292-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/696,668
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/427,154
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3816
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-696-668-2

Query Match 89.4%; Score 3393.4; DB 4; Length 3816;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db        422 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 481

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
Db        482 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 541

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db        542 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 601

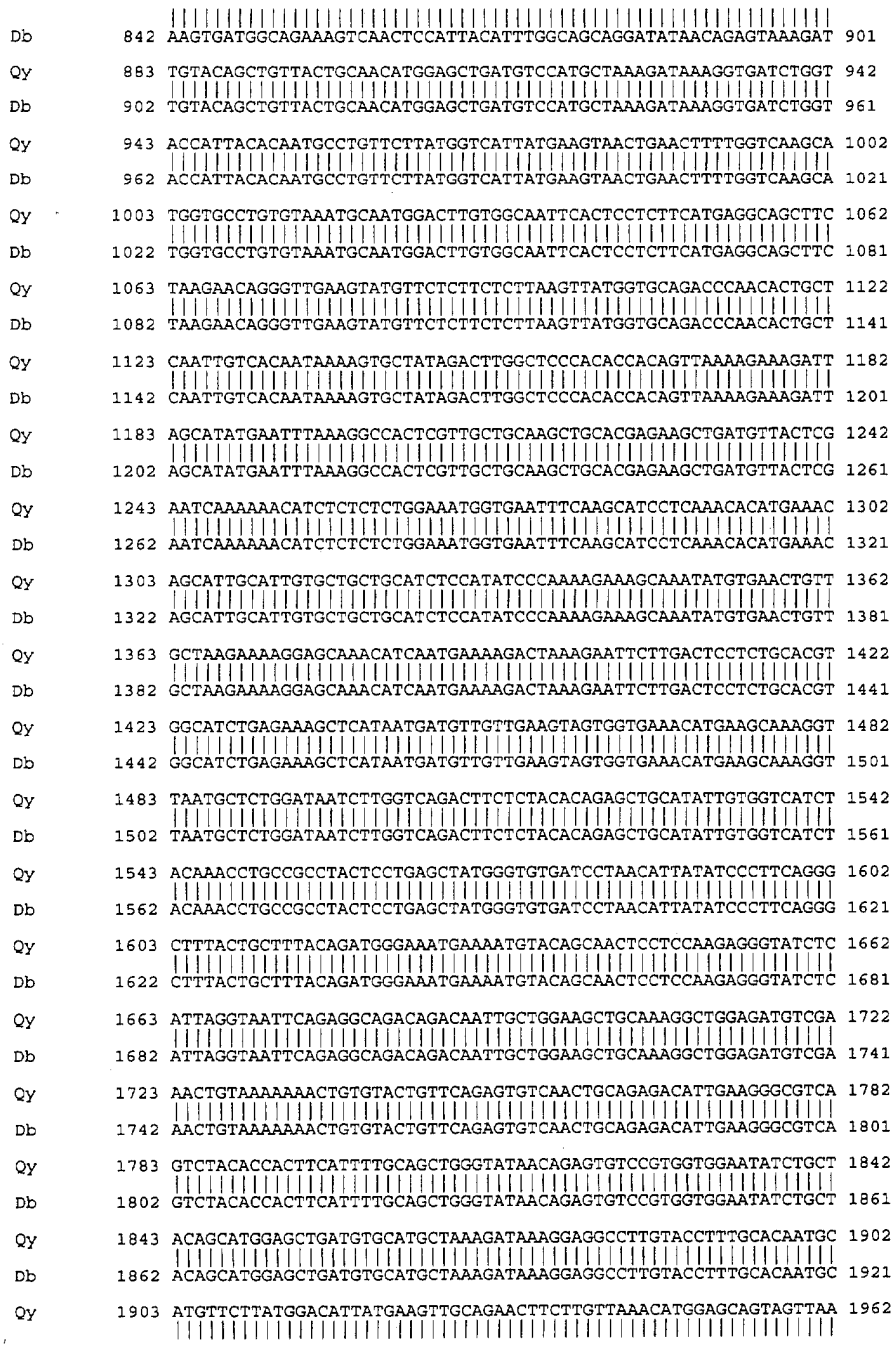
Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db        602 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 661

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db        662 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 721

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db        722 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 781

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db        782 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 841

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
```


Db 842 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 901
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db 902 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 961
Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 962 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1021
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 1022 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1081
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 1082 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1141
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 1142 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1201
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 1202 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1261
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1262 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1321
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1322 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1381
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1382 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1441
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1442 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1501
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1502 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1561
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1562 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1621
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1622 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1681
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1682 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1741
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1742 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1801
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1802 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1861
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1862 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1921
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1922 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1981
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1982 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2041
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2082
Db 2042 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAA 2101
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2102 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2161
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2162 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2221
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2222 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2281
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2282 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2341
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2342 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2401
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2402 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2461
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2462 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2521
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2522 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2581
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2582 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2641
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2642 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2701
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2702 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2761
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2762 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2821
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2822 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2881
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2882 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2941
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 2942 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3001
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3002 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3061

```
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3121

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3181

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3241

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3301

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3361

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3421

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3481

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3541

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3601

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3661

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3721

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3781

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
             ||||||||||||||||||||||||||||||||||||
Db       3782 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3816
```

RESULT 3
US-09-972-115A-5
; Sequence 5, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5
; LENGTH: 4275
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (284)..(3781)
; OTHER INFORMATION:
US-09-972-115A-5

Query Match 89.2%; Score 3388.8; DB 4; Length 4275;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        481 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 540

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 600

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 660

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 720

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 780

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 840

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 900

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 960

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1020

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1080

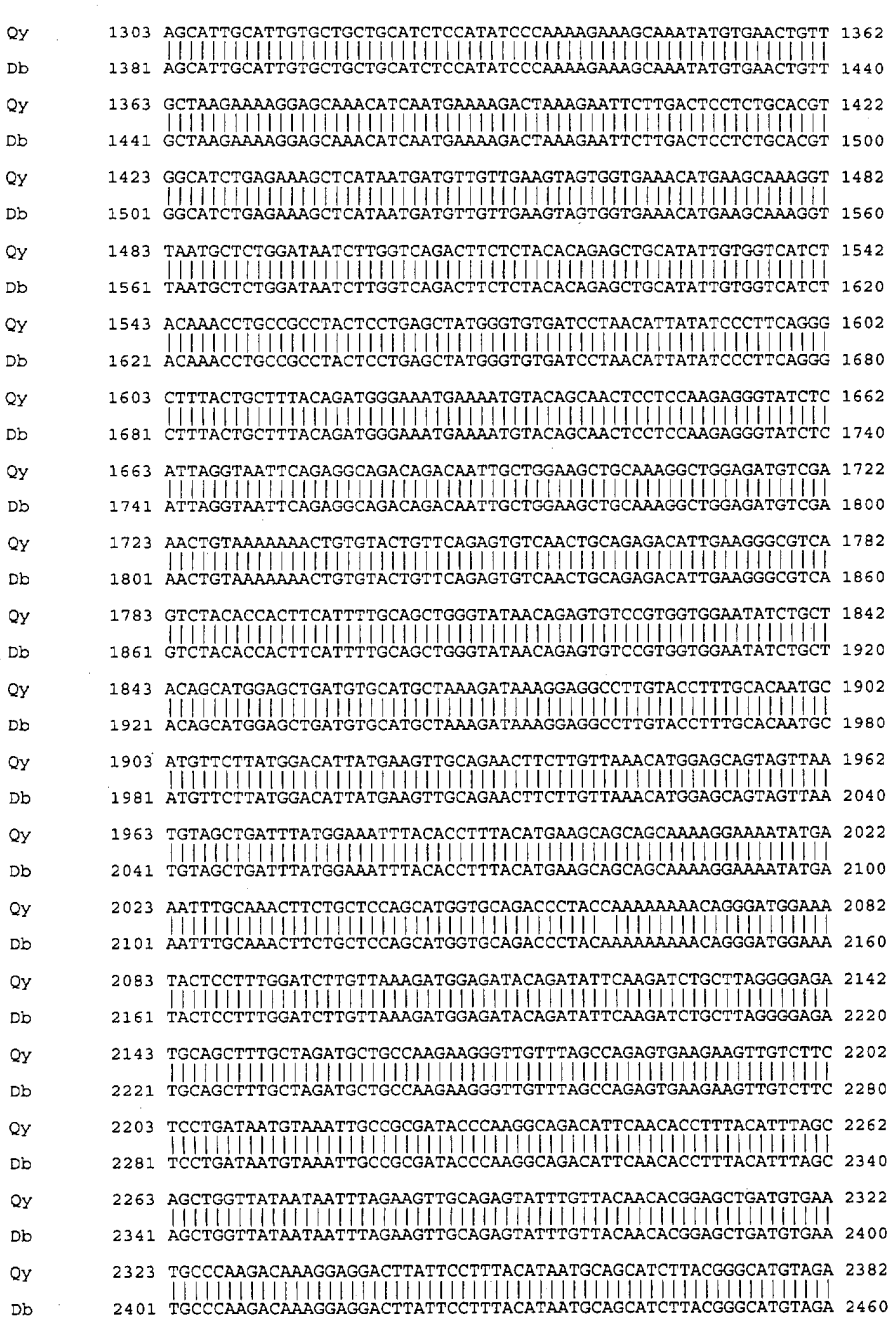
Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1140

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1200

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1260

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1320

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1380
```


Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1381 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1440
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1441 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1500
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1501 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1560
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1561 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1620
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1621 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1680
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1681 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1740
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1741 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1800
Qy 1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1801 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1860
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1861 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1920
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1921 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1980
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1981 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 2040
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 2041 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2100
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 2101 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 2160
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 2161 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2220
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 2221 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2280
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 2281 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2340
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2341 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2400
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2401 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2460
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2461 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2520
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2521 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2580
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2581 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2640
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2641 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2700
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2701 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2760
Qy 2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db 2761 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2820
Qy 2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db 2821 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2880
Qy 2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db 2881 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2940
Qy 2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db 2941 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 3000
Qy 2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db 3001 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 3060
Qy 2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db 3061 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3120
Qy 3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db 3121 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3180
Qy 3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db 3181 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3240
Qy 3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db 3241 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3300
Qy 3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db 3301 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3360
Qy 3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db 3361 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3420
Qy 3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db 3421 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3480
Qy 3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db 3481 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3540
Qy 3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522

```
Db       3541 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3600

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582

Db       3601 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3660

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642

Db       3661 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3720

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702

Db       3721 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3780

Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762

Db       3781 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3840

Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794

Db       3841 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3872
```

RESULT 4
US-09-427-154-1
; Sequence 1, Application US/09427154
; Patent No. 6589725
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Chan, Eva
; APPLICANT: Xu, Xiang
; APPLICANT: Huang, Betty
; TITLE OF INVENTION: TANKYRASE H, COMPOSTIONS INVOLVED IN THE CELL CYCLE AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: A-68292-DJB/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/427,154
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-427-154-1

Query Match 89.2%; Score 3386; DB 4; Length 3394;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3389; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        404 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGCA 463

Db          1 GGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGCA 60

Qy        464 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 523

Db         61 CGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGTA 120

Qy        524 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 583

Db        121 GTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATACT 180

Qy        584 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 643

Db        181 CCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCAT 240

Qy        644 GGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 703

Db        241 GGAGCTGAGCCAACCATCCTAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCCA 300

Qy        704 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 763

Db        301 TCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAGG 360

Qy        764 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 823

Db        361 AGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGCA 420

Qy        824 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 883

Db        421 AGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGATT 480

Qy        884 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 943

Db        481 GTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTA 540

Qy        944 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 1003

Db        541 CCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCAT 600

Qy       1004 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 1063

Db        601 GGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCT 660

Qy       1064 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 1123

Db        661 AAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTC 720

Qy       1124 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 1183

Db        721 AATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTA 780

Qy       1184 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 1243

Db        781 GCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGA 840

Qy       1244 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 1303

Db        841 ATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACA 900

Qy       1304 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 1363

Db        901 GCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTTG 960

Qy       1364 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1423

Db        961 CTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTG 1020

Qy       1424 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1483

Db       1021 GCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTT 1080

Qy       1484 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1543

Db       1081 AATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTA 1140

Qy       1544 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1603

Db       1141 CAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGC 1200

Qy       1604 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1663

Db       1201 TTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCA 1260

Qy       1664 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1723

Db       1261 TTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAA 1320

Qy       1724 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1783

Db       1321 ACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAG 1380

Qy       1784 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1843

Db       1381 TCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTA 1440
```

```
Qy       1844 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1903
Db       1441 CAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCA 1500

Qy       1904 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1963
Db       1501 TGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAAT 1560

Qy       1964 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 2023
Db       1561 GTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAA 1620

Qy       2024 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 2083
Db       1621 ATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAAT 1680

Qy       2084 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGAT 2143
Db       1681 ACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCATTATCTGCTTAGGGGAGAT 1740

Qy       2144 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 2203
Db       1741 GCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCT 1800

Qy       2204 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 2263
Db       1801 CCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGCA 1860

Qy       2264 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 2323
Db       1861 GCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAAT 1920

Qy       2324 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 2383
Db       1921 GCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGAT 1980

Qy       2384 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2443
Db       1981 GTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTTC 2040

Qy       2444 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2503
Db       2041 ACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGCC 2100

Qy       2504 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2563
Db       2101 CATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTCA 2160

Qy       2564 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2623
Db       2161 GCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTGT 2220

Qy       2624 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2683
Db       2221 TACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTCT 2280

Qy       2684 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2743
Db       2281 TCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGG 2340

Qy       2744 AGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2803
Db       2341 AGTTTTTCAGAACTGTCTTCATTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTG 2400

Qy       2804 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2863
Db       2401 GAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGA 2460

Qy       2864 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2923
Db       2461 CTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAG 2520

Qy       2924 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2983
Db       2521 ATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTA 2580

Qy       2984 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 3043
Db       2581 ATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTG 2640

Qy       3044 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 3103
Db       2641 AACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAG 2700

Qy       3104 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 3163
Db       2701 TCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGT 2760

Qy       3164 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 3223
Db       2761 GGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGG 2820

Qy       3224 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 3283
Db       2821 GAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAA 2880

Qy       3284 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 3343
Db       2881 CGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAA 2940

Qy       3344 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3403
Db       2941 AGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCC 3000

Qy       3404 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3463
Db       3001 AAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGAC 3060

Qy       3464 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3523
Db       3061 AGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCT 3120

Qy       3524 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3583
Db       3121 TTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACT 3180

Qy       3584 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3643
Db       3181 GGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAG 3240

Qy       3644 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3703
Db       3241 GCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGGA 3300

Qy       3704 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3763
Db       3301 TAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGGC 3360

Qy       3764 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3797
Db       3361 CTCTACGTTTTACTCCTTTGCTGAAAAAAAAAAA 3394
```

RESULT 5
US-09-350-982C-3
; Sequence 3, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
; APPLICANT: Berthelsen, Jens
; APPLICANT: Toma, Salvatore
; APPLICANT: Isacchi, Antonella
; TITLE OF INVENTION: Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Rela
; TITLE OF INVENTION: Same
; FILE REFERENCE: PHRM-0043
; CURRENT APPLICATION NUMBER: US/09/350,982C
; CURRENT FILING DATE: 1999-07-09

```
;  NUMBER OF SEQ ID NOS:   10
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 4512
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: PCR Primers
;   NAME/KEY: misc_feature
;   LOCATION: (1124)..(1124)
;   OTHER INFORMATION: n is any nucleic acid
;   NAME/KEY: misc_feature
;   LOCATION: (2672)..(2672)
;   OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-3

  Query Match             89.1%;  Score 3382.8;  DB 4;  Length 4512;
  Best Local Similarity   99.7%;  Pred. No. 0;
  Matches 3381;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 279

Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 339

Qy        523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 399

Qy        583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 459

Qy        643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 519

Qy        703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 579

Qy        763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 639

Qy        823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 699

Qy        883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 759

Qy        943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 819

Qy       1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 879

Qy       1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 939

Qy       1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         940 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 999

Qy       1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1000 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1059

Qy       1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1060 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1119

Qy       1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1120 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1179

Qy       1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        1180 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1239

Qy       1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
              ||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db        1240 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1299

Qy       1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1300 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1359

Qy       1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1360 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1419

Qy       1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1420 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1479

Qy       1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1480 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1539

Qy       1723 AACTGTAAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1540 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1599

Qy       1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1600 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1659

Qy       1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
              ||||||||||||||||||||||||||||||||::|:||||||||||||||||||||||||
Db        1660 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1719

Qy       1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1720 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1779

Qy       1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1780 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1839

Qy       2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1840 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1899

Qy       2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1900 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1959

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1960 TGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2019

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2020 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2079

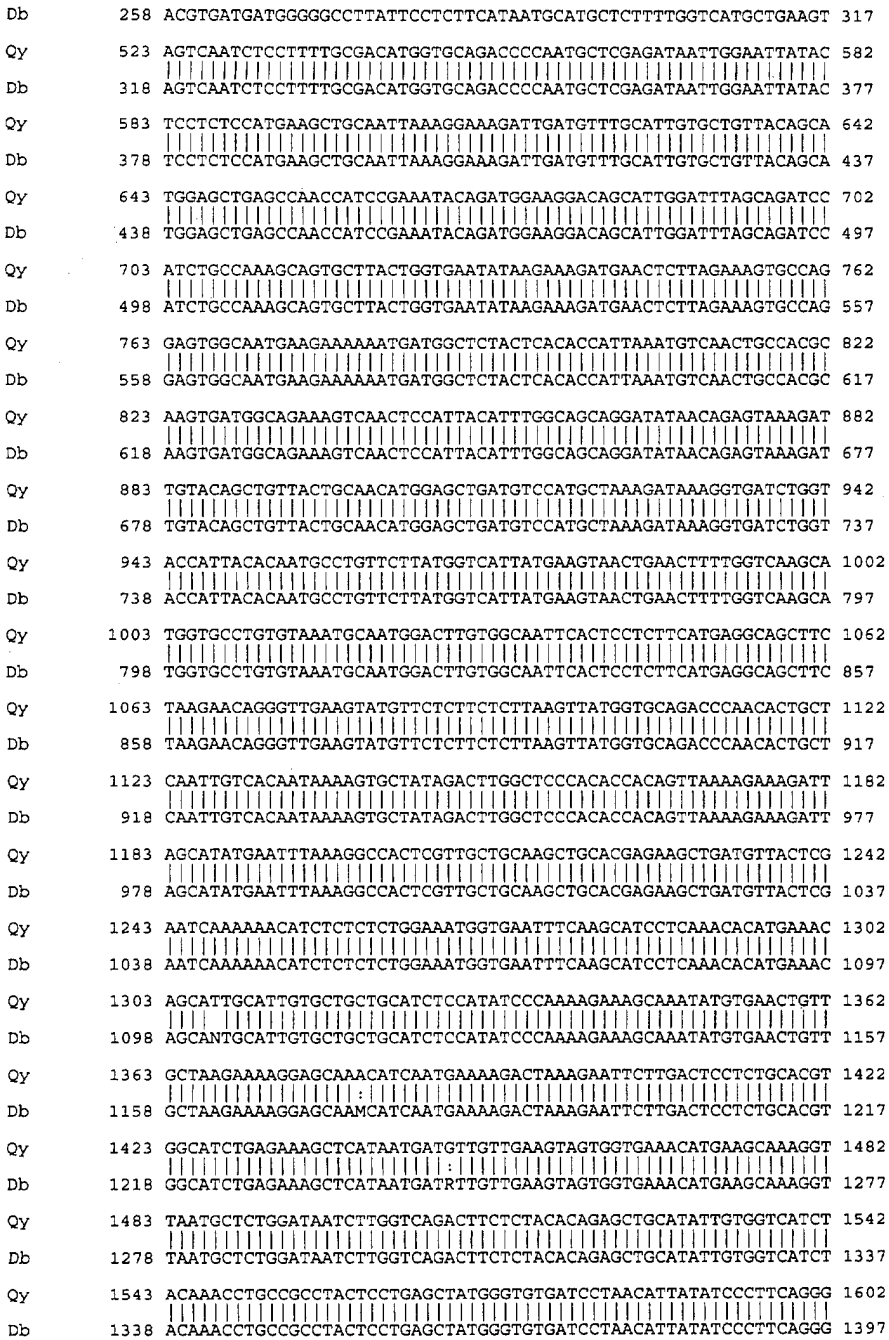
Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2080 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2139
Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       2140 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2199
Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       2200 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2259
Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       2260 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGKTGCTAGC 2319
Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       2320 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2379
Qy       2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db       2380 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2439
Qy       2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db       2440 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2499
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
Db       2500 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2559
Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
Db       2560 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2619
Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
Db       2620 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGNATCTTGG 2679
Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
Db       2680 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2739
Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
Db       2740 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2799
Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
Db       2800 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2859
Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
Db       2860 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2919
Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
Db       2920 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2979
Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
Db       2980 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3039
Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
Db       3040 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3099
Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
Db       3100 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3159
Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
Db       3160 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3219
Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
Db       3220 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3279
Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
Db       3280 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3339
Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
Db       3340 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3399
Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
Db       3400 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3459
Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
Db       3460 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3519
Qy       3703 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3762
Db       3520 ATAAATAGTTATTTTAAGAAACTAATTCCACTGAACCTAAAATCATCAAAGCAGCAGTGG 3579
Qy       3763 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
Db       3580 CCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3611
```

```
RESULT 6
US-09-350-982C-4
; Sequence 4, Application US/09350982C
; Patent No. 6455290
; GENERAL INFORMATION:
;  APPLICANT:  Berthelsen, Jens
;  APPLICANT:  Toma, Salvatore
;  APPLICANT:  Isacchi, Antonella
;  TITLE OF INVENTION:  Tankyrase Homolog Protein(THP), Nucleic Acids, and Methods Rela
;  TITLE OF INVENTION:  Same
;  FILE REFERENCE:  PHRM-0043
;  CURRENT APPLICATION NUMBER: US/09/350,982C
;  CURRENT FILING DATE:  1999-07-09
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 4
;  LENGTH: 3498
;  TYPE: DNA
;  ORGANISM: Artificial
;  FEATURE:
;  OTHER INFORMATION: PCR Primers
;  NAME/KEY: CDS
;  LOCATION: (1)..(3498)
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: Xaa is any amino acid
;  NAME/KEY: misc_feature
;  LOCATION: (1102)..(1102)
;  OTHER INFORMATION: n is any nucleic acid
;  NAME/KEY: misc_feature
;  LOCATION: (2650)..(2650)
;  OTHER INFORMATION: n is any nucleic acid
US-09-350-982C-4

  Query Match           86.7%;  Score 3291.8;  DB 4;  Length 3498;
  Best Local Similarity  99.7%;  Pred. No. 0;
  Matches 3290;  Conservative    6;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
Db        198 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAATGTCCAAGC 257
Qy        463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
```


Db 258 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 317
Qy 523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
Db 318 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 377
Qy 583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
Db 378 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 437
Qy 643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
Db 438 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 497
Qy 703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
Db 498 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 557
Qy 763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
Db 558 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 617
Qy 823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
Db 618 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 677
Qy 883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
Db 678 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 737
Qy 943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
Db 738 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 797
Qy 1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
Db 798 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 857
Qy 1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
Db 858 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 917
Qy 1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
Db 918 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 977
Qy 1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
Db 978 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1037
Qy 1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
Db 1038 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1097
Qy 1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
Db 1098 AGCANTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1157
Qy 1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
Db 1158 GCTAAGAAAAGGAGCAAMCATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1217
Qy 1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db 1218 GGCATCTGAGAAAGCTCATAATGATRTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1277
Qy 1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
Db 1278 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1337
Qy 1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
Db 1338 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1397
Qy 1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
Db 1398 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1457
Qy 1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
Db 1458 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1517
Qy 1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
Db 1518 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1577
Qy 1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
Db 1578 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1637
Qy 1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
Db 1638 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGRRGSCCTTGTACCTTTGCACAATGC 1697
Qy 1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
Db 1698 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1757
Qy 1963 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 2022
Db 1758 TGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGA 1817
Qy 2023 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAA 2082
Db 1818 AATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACAAAAAAAAACAGGGATGGAAA 1877
Qy 2083 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 2142
Db 1878 TACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGA 1937
Qy 2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db 1938 TGCAGCTTTGCTAGATGCTGCCAAGAAGAGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 1997
Qy 2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db 1998 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2057
Qy 2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db 2058 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2117
Qy 2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db 2118 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2177
Qy 2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db 2178 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2237
Qy 2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db 2238 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGKTGCTAGC 2297
Qy 2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db 2298 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2357
Qy 2563 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2622
Db 2358 AGCGGATGATGTCAGCGCTCTTCTGACAGCAGCCATGCCCCCATCTGCTCTGCCCTCTTG 2417
Qy 2623 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2682
Db 2418 TTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGCAGATGCTCTCTC 2477

```
Qy       2683 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2478 TTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGG 2537

Qy       2743 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2538 GAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTT 2597

Qy       2803 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGG 2862
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       2598 GGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGNATCTTGG 2657

Qy       2863 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2922
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2658 ACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGA 2717

Qy       2923 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2982
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2718 GATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACT 2777

Qy       2983 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 3042
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2778 AATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTT 2837

Qy       3043 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 3102
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2838 GAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCA 2897

Qy       3103 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 3162
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2898 GTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGG 2957

Qy       3163 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3222
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2958 TGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATG 3017

Qy       3223 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3282
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3018 GGAAAGATACACTCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGA 3077

Qy       3283 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3078 ACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGA 3137

Qy       3343 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3138 AAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTC 3197

Qy       3403 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3198 CAAAAGCAATCAATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGA 3257

Qy       3463 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3258 CAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTC 3317

Qy       3523 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3318 TTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCAC 3377

Qy       3583 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3378 TGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACA 3437

Qy       3643 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3438 GGCTTATCCTGAGTATTTAATTACTTACCAGATTATGAGGCCTGAAGGTATGGTCGATGG 3497

Qy       3703 A 3703
              |
Db       3498 A 3498
```

```
RESULT 7
US-09-972-115A-1
; Sequence 1, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(3999)
;   OTHER INFORMATION:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(586)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (596)..(597)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (621)..(621)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (626)..(626)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (629)..(629)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (638)..(638)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (693)..(694)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (724)..(724)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (730)..(730)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (764)..(764)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (791)..(791)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2148)..(2148)
```

```
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2177)..(2177)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2382)..(2382)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2398)..(2398)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2466)..(2466)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2468)..(2468)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2493)..(2493)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (2536)..(2860)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (3163)..(3163)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
;   NAME/KEY: misc_feature
;   LOCATION: (3488)..(3499)
;   OTHER INFORMATION: Sequences not determined at this position .  Unknown residue put
;   OTHER INFORMATION: in to provide alignment.
US-09-972-115A-1

  Query Match           78.2%;  Score 2969.6;  DB 4;  Length 4493;
  Best Local Similarity  88.6%;  Pred. No. 0;
  Matches 3028;  Conservative    1;  Mismatches  363;  Indels   27;  Gaps    2;

Qy       403 AGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAGAATGGTGCAAGTGTCCAAGC 462
            |||||||||||||||| || |  || |||||||||||| |||||||||| | ||||| |
Db       672 AGGTTTTGGGCGGAAAAACCTNNTTAAATATTTGCTTCAAAATGGTGCAAATNTCCAANC 731

Qy       463 ACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGCTCTTTTGGTCATGCTGAAGT 522
            || | || |||||||||||||||||||| |||||||||||||||||||| ||
Db       732 ACTTTATAATGGGGGCCTTATTCCTCTTCATANTGCATGCTCTTTTGGTCATGCTAAAAN 791

Qy       523 AGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       792 TATCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCTCGAGATAATTGGAATTATAC 851

Qy       583 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       852 TCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTTTGCATTGTGCTGTTACAGCA 911

Qy       643 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       912 TGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACAGCATTGGATTTAGCAGATCC 971

Qy       703 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       972 ATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGATGAACTCTTAGAAAGTGCCAG 1031

Qy       763 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 822
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1032 GAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCATTAAATGTCAACTGCCACGC 1091

Qy       823 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 882
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1092 AAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCAGGATATAACAGAGTAAAGAT 1151

Qy       883 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 942
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1152 TGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGT 1211

Qy       943 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1002
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1212 ACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCA 1271

Qy      1003 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1062
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1272 TGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTC 1331

Qy      1063 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1332 TAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCT 1391

Qy      1123 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1392 CAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATT 1451

Qy      1183 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1452 AGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCG 1511

Qy      1243 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1512 AATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAAC 1571

Qy      1303 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1362
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1572 AGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAACTGTT 1631

Qy      1363 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1422
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1632 GCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGT 1691

Qy      1423 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1692 GGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGT 1751

Qy      1483 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1752 TAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCT 1811

Qy      1543 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1812 ACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGG 1871

Qy      1603 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1872 CTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTC 1931

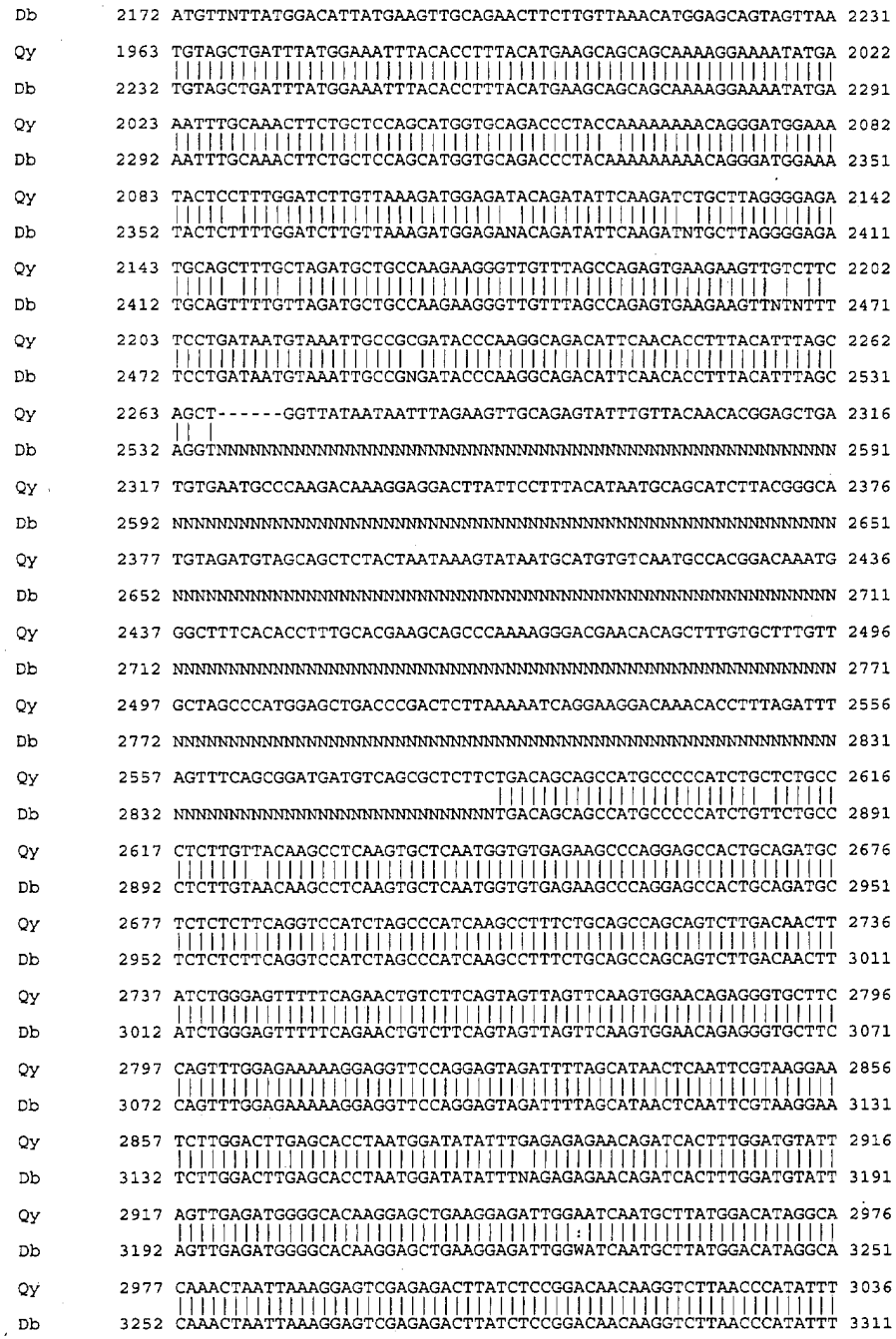
Qy      1663 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1932 ATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGA 1991

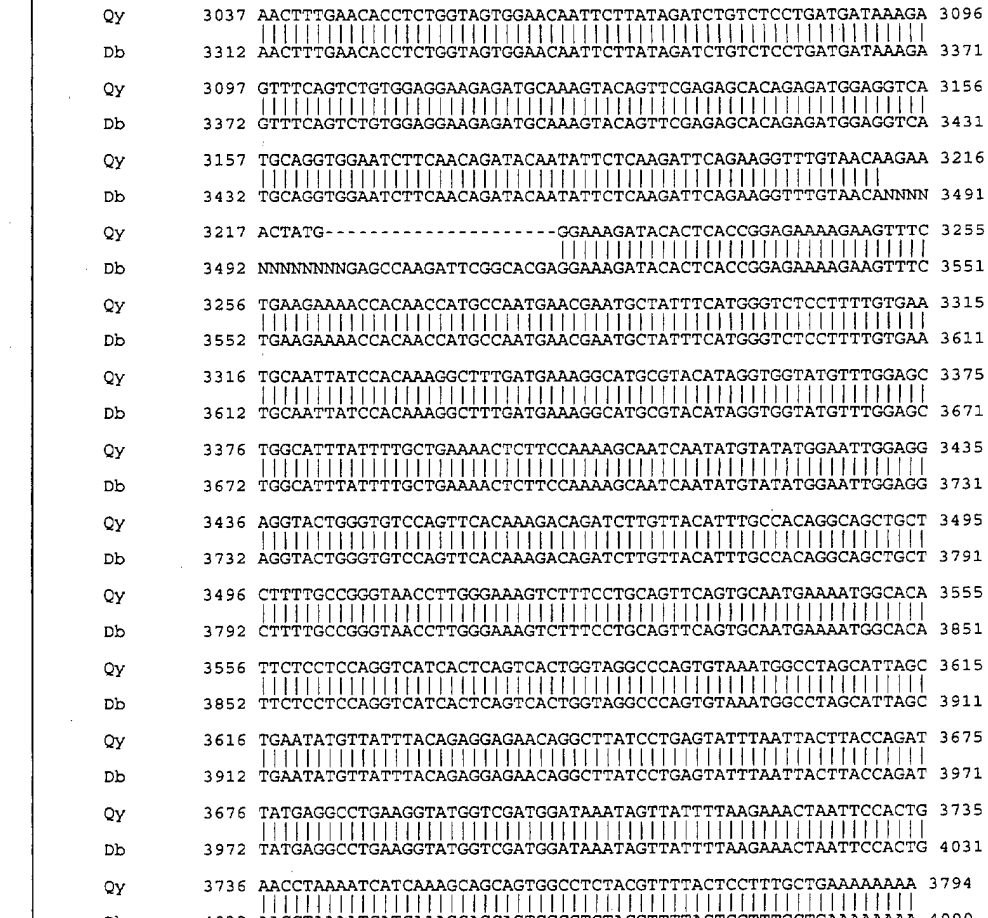
Qy      1723 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 1782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1992 AACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCA 2051

Qy      1783 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 1842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2052 GTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCT 2111

Qy      1843 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGC 1902
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db      2112 ACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGNGGCCTTGTACCTTTGCACAATGC 2171

Qy      1903 ATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAA 1962
             ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

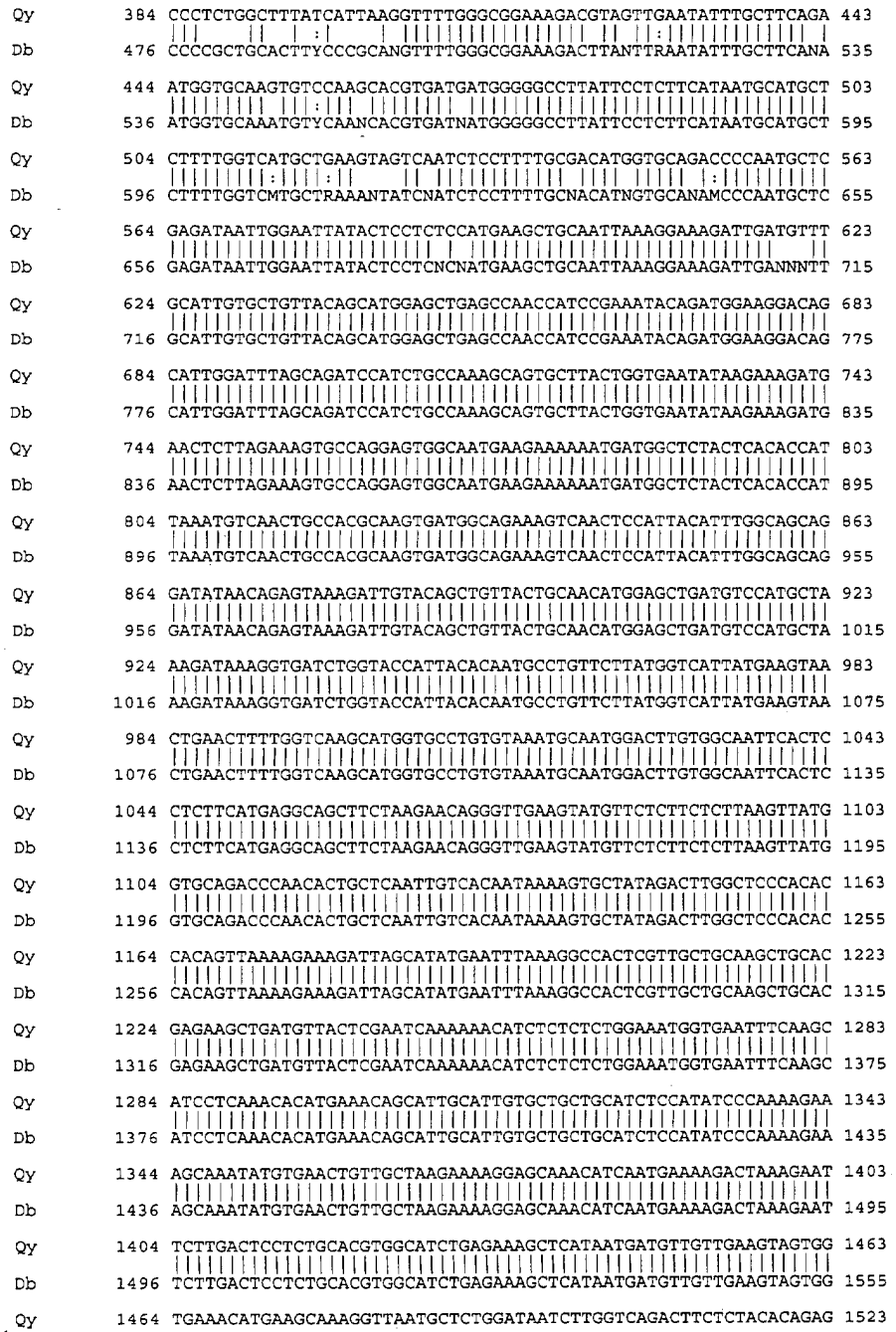


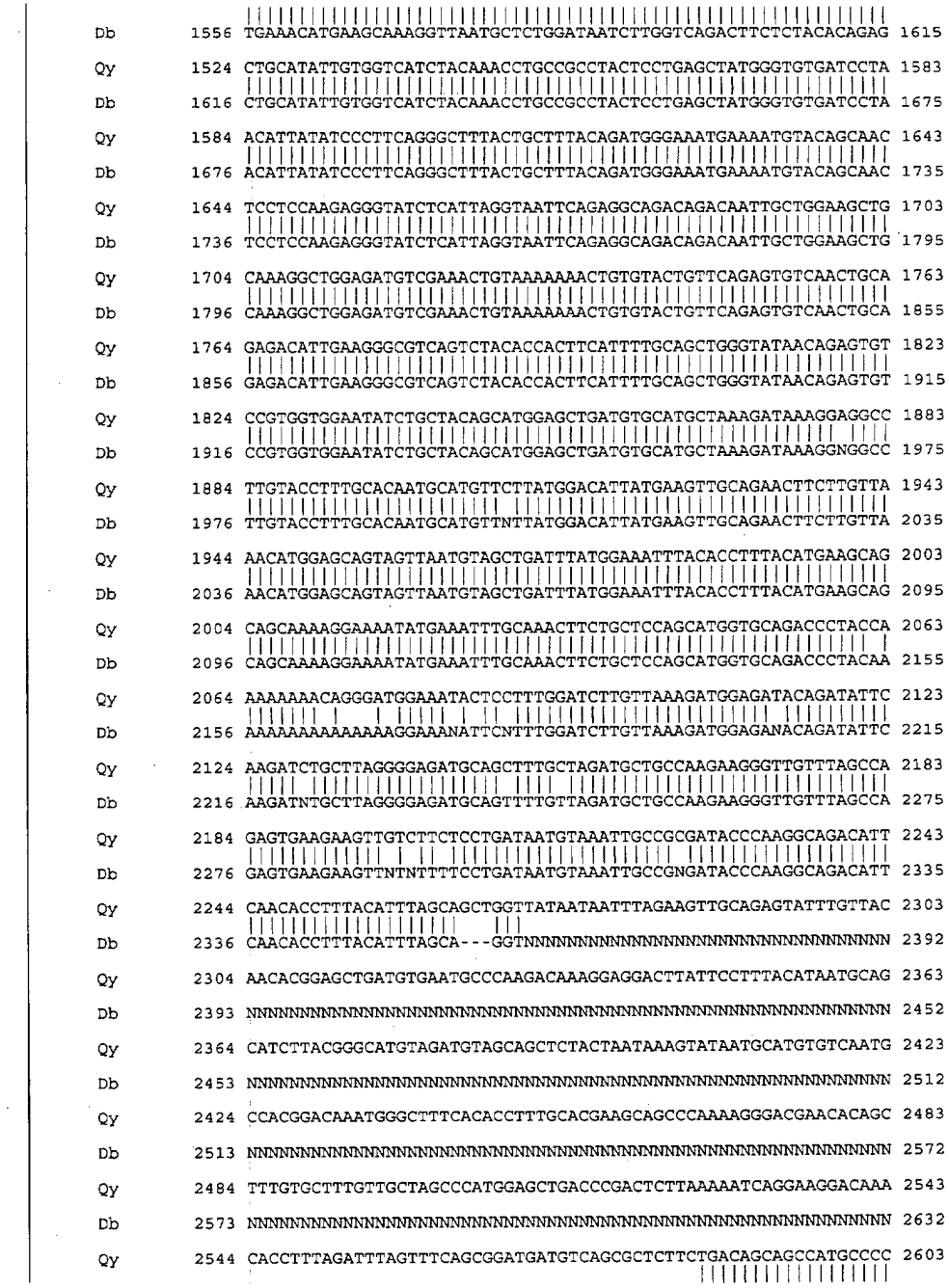


```
RESULT 8
US-09-972-115A-3
; Sequence 3, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
```

; SEQ ID NO 3
; LENGTH: 4297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3801)
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (125)..(125)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (293)..(293)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (328)..(328)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (341)..(341)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (378)..(378)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (415)..(415)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (459)..(459)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (496)..(496)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (517)..(517)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (534)..(534)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (553)..(553)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (562)..(562)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (614)..(614)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (619)..(619)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (632)..(632)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (637)..(637)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (643)..(643)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (680)..(680)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (682)..(682)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (711)..(713)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (1971)..(1971)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2000)..(2000)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2175)..(2175)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2180)..(2180)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2205)..(2205)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2221)..(2221)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2289)..(2289)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2316)..(2316)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2359)..(2674)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2977)..(2977)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
; NAME/KEY: misc_feature
; LOCATION: (2291)..(2291)
; OTHER INFORMATION: Sequences not determined at this position . Unknown residue put
; OTHER INFORMATION: in to provide alignment.
US-09-972-115A-3

Query Match 78.2%; Score 2968.2; DB 4; Length 4297;
Best Local Similarity 88.3%; Pred. No. 0;
Matches 3020; Conservative 7; Mismatches 381; Indels 12; Gaps 2;





```
Db       2633 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGACAGCAGCCATGCCCC 2692

Qy       2604 CATCTGCTCTGCCCTCTTGTTACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAG 2663
              |||||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       2693 CATCTGTTCTGCCCTCTTGTAACAAGCCTCAAGTGCTCAATGGTGTGAGAAGCCCAGGAG 2752

Qy       2664 CCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCA 2723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 CCACTGCAGATGCTCTCTCTTCAGGTCCATCTAGCCCATCAAGCCTTTCTGCAGCCAGCA 2812

Qy       2724 GTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAA 2783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTCTTCAGTAGTTAGTTCAAGTGGAA 2872

Qy       2784 CAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTC 2843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 CAGAGGGTGCTTCCAGTTTGGAGAAAAAGGAGGTTCCAGGAGTAGATTTTAGCATAACTC 2932

Qy       2844 AATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAACAGATCA 2903
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       2933 AATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTNAGAGAGAACAGATCA 2992

Qy       2904 CTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATCAATGCTT 2963
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db       2993 CTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGWATCAATGCTT 3052

Qy       2964 ATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAACAAGGTC 3023
              |||||||||||||||||||||||| |  ||||||||||||||||||||||||||||||
Db       3053 ATGGACATAGGCACAAACTAATTAAAAGTTTCGAGAGACTTATCTCCGGACAACAAGGTC 3112

Qy       3024 TTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTC 3083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 TTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGATCTGTCTC 3172

Qy       3084 CTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACA 3143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGAGAGCACA 3232

Qy       3144 GAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGG 3203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATTCAGAAGG 3292

Qy       3204 TTTGTAAC---------AAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3254
              |||||||          |    |    ||||||||||||||||||||||||||||||||
Db       3293 TTTGTAACAGAGCCAAGATTCGGCACGAGGAAAGATACACTCACCGGAGAAAAGAAGTTT 3352

Qy       3255 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 CTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGA 3412

Qy       3315 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 ATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAG 3472

Qy       3375 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 CTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAG 3532

Qy       3435 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3494
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGC 3592

Qy       3495 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 3554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 TCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCAC 3652

Qy       3555 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAG 3712

Qy       3615 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 CTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGA 3772

Qy       3675 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3773 TTATGAGGCCTGAAGGTATGGTCGATGGATAAATAGTTATTTTAAGAAACTAATTCCACT 3832

Qy       3735 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 GAACCTAAAATCATCAAAGCAGCAGTGGCCTCTACGTTTTACTCCTTTGCTGAAAAAAAA 3892
```

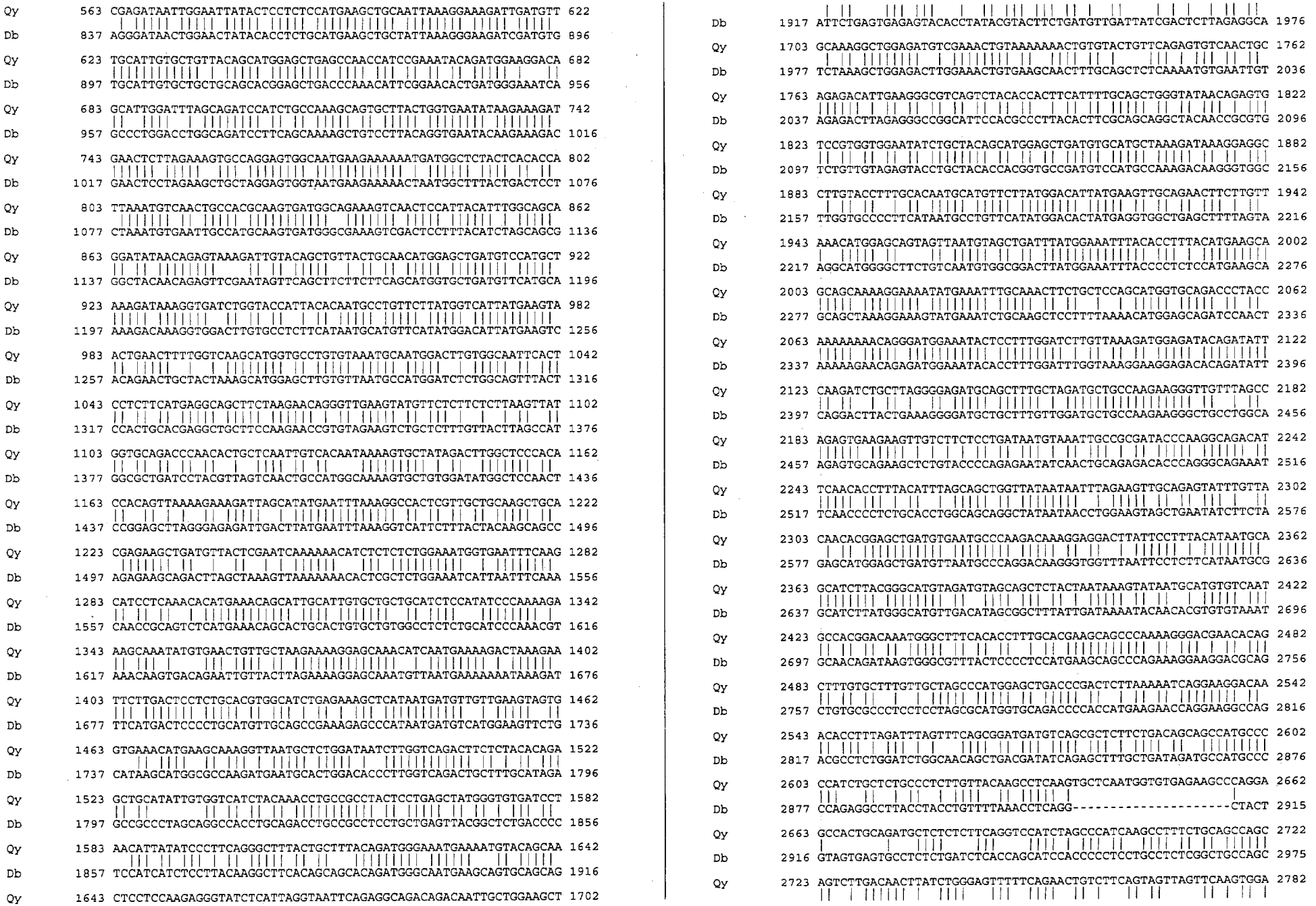
RESULT 9
US-09-196-387-1
; Sequence 1, Application US/09196387
; Patent No. 6277613
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4134 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-196-387-1

Query Match 47.8%; Score 1813.4; DB 3; Length 4134;
Best Local Similarity 72.3%; Pred. No. 0;
Matches 2395; Conservative 0; Mismatches 891; Indels 27; Gaps 2;

```
Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
              || | |  ||  |    | ||||||||  |||| || || || ||| | || || |||
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
              |  |||||  |||||| || |||||||||||  || || || || ||||||||||| ||
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
              |||||||| |||||||| || || | |||  | ||| | || || || || || ||||||
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836
```



```
Db       2976 AGCATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGCCGTAGGAGGAGCCTCCAATGCA 3035

Qy       2783 ACAGAGGGTGCTTCCAGTTTGGAGAAAAAG------GAGGTTCCAGGAGTAGATTTTAGC 2836

Db       3036 GGGGATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGAAGTTGCTGGTCTTGACATGAAT 3095

Qy       2837 ATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCTAATGGATATATTTGAGAGAGAA 2896

Db       3096 ATCAGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCTTCGGGATATCTTTGAAACAGAA 3155

Qy       2897 CAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAAGGAGCTGAAGGAGATTGGAATC 2956

Db       3156 CAGATTACACTAGATGTGTTGGCTGATATGGGTCATGAAGAGTTGAAAGAAATAGGCATC 3215

Qy       2957 AATGCTTATGGACATAGGCACAAACTAATTAAAGGAGTCGAGAGACTTATCTCCGGACAA 3016

Db       3216 AATGCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAA 3275

Qy       3017 CAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 3076

Db       3276 CAAGGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGAT 3335

Qy       3077 CTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGA 3136

Db       3336 CTTGCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGA 3395

Qy       3137 GAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATT 3196

Db       3396 GAACACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATT 3455

Qy       3197 CAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCT 3256

Db       3456 CAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCT 3515

Qy       3257 GAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAAT 3316

Db       3516 GAGGAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAAT 3575

Qy       3317 GCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCT 3376

Db       3576 GCCATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCC 3635

Qy       3377 GGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA 3436

Db       3636 GGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGA 3695

Qy       3437 GGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTC 3496

Db       3696 GGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTC 3755

Qy       3497 TTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACAT 3556

Db       3756 TTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCAC 3815

Qy       3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616

Db       3816 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3875

Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676

Db       3876 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3935

Qy       3677 ATGAGGCCTGAAG 3689

Db       3936 ATGAAGCCAGAAG 3948


RESULT 10
US-09-841-835-1
; Sequence 1, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4134 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-841-835-1

  Query Match             47.8%;  Score 1813.4;  DB 4;  Length 4134;
  Best Local Similarity   72.3%;  Pred. No. 0;
  Matches 2395;  Conservative    0;  Mismatches  891;  Indels   27;  Gaps    2;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682

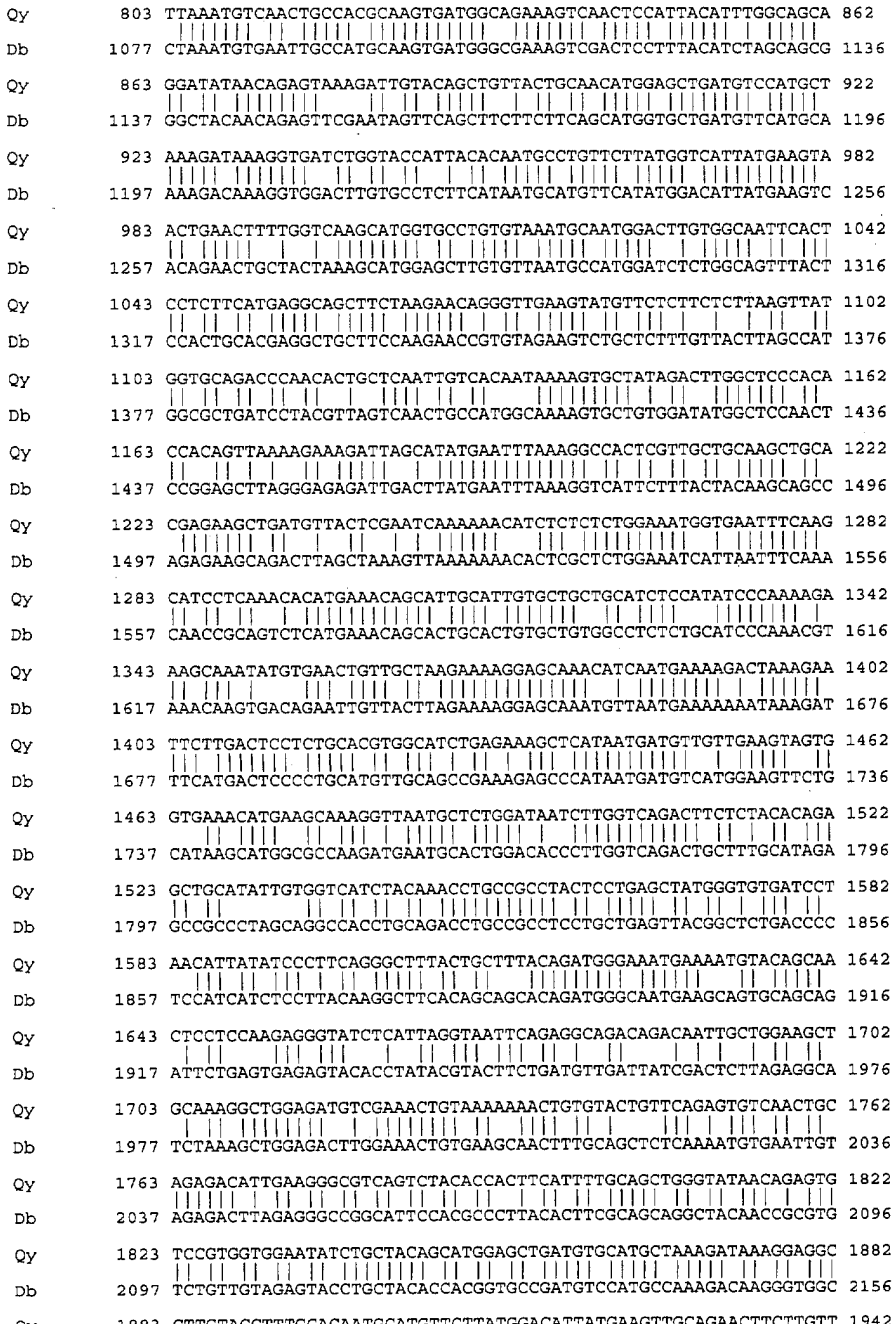
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

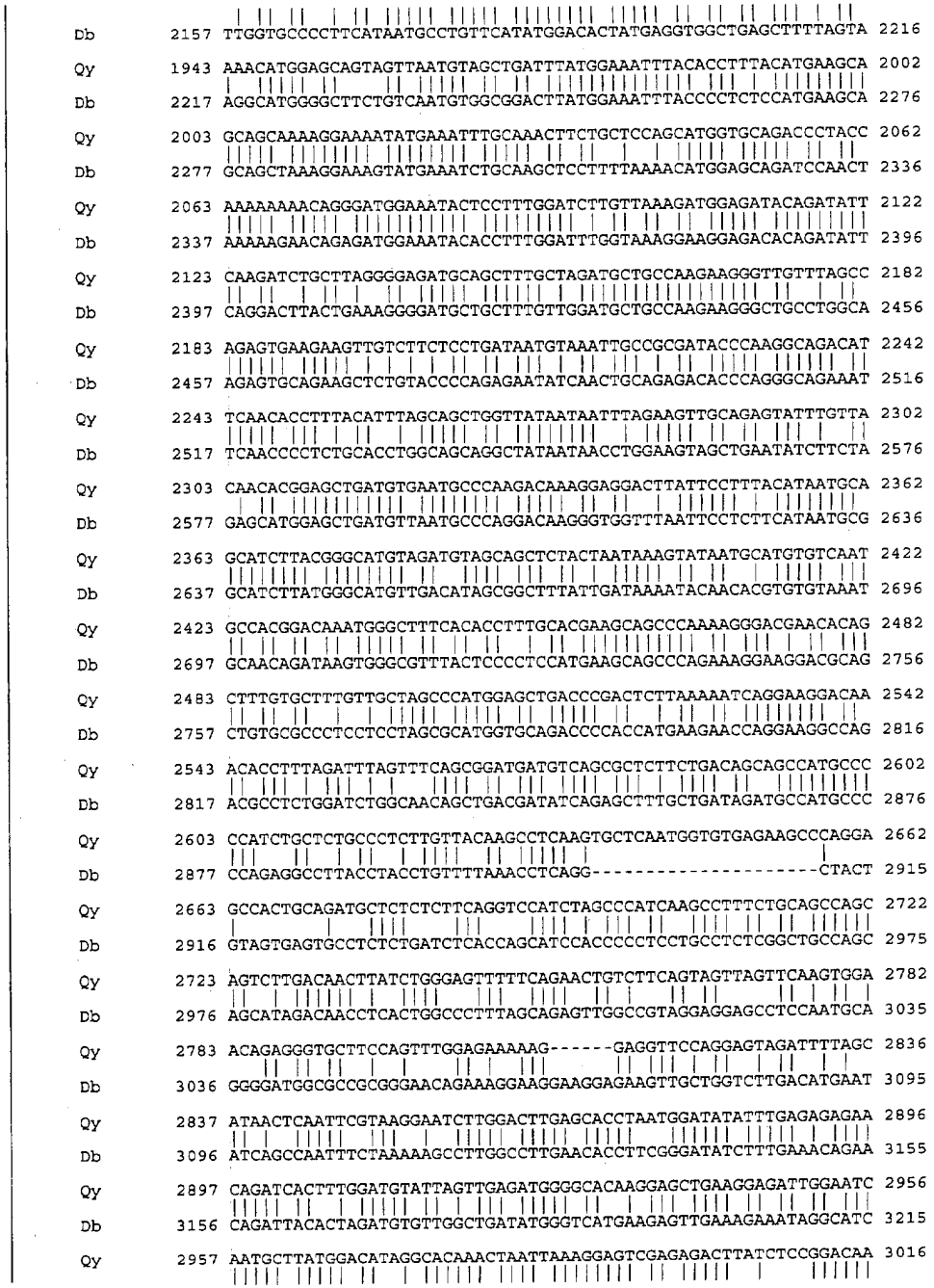
Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742

Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802

Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076
```





```
Db       3216 AATGCATATGGGCACCGCCACAAATTAATCAAAGGAGTAGAAAGACTCTTAGGTGGACAA 3275

Qy       3017 CAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGGTAGTGGAACAATTCTTATAGAT 3076

Db       3276 CAAGGCACCAATCCTTATTTGACTTTTCACTGTGTTAATCAGGGAACGATTTTGCTGGAT 3335

Qy       3077 CTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGAAGAGATGCAAAGTACAGTTCGA 3136

Db       3336 CTTGCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGAAGAGATGCAAAGTACTATTCGA 3395

Qy       3137 GAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAACAGATACAATATTCTCAAGATT 3196

Db       3396 GAACACAGAGATGGTGGTAATGCTGGCGGCATCTTCAACAGATACAATGTCATTCGAATT 3455

Qy       3197 CAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACACTCACCGGAGAAAAGAAGTTTCT 3256

Db       3456 CAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCT 3515

Qy       3257 GAAGAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAAT 3316

Db       3516 GAGGAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAAT 3575

Qy       3317 GCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCT 3376

Db       3576 GCCATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCC 3635

Qy       3377 GGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGA 3436

Db       3636 GGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGA 3695

Qy       3437 GGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTC 3496

Db       3696 GGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTC 3755

Qy       3497 TTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACAT 3556

Db       3756 TTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCAC 3815

Qy       3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616

Db       3816 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3875

Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676

Db       3876 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3935

Qy       3677 ATGAGGCCTGAAG 3689

Db       3936 ATGAAGCCAGAAG 3948
```

```
RESULT 11
US-09-972-115A-7
; Sequence 7, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 4134
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-7

  Query Match             47.8%;  Score 1813.4;  DB 4;  Length 4134;
  Best Local Similarity   72.3%;  Pred. No. 0;
  Matches 2395;  Conservative    0;  Mismatches  891;  Indels   27;  Gaps    2;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682

Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742

Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802

Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862

Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922

Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982

Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042

Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102

Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162

Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222

Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

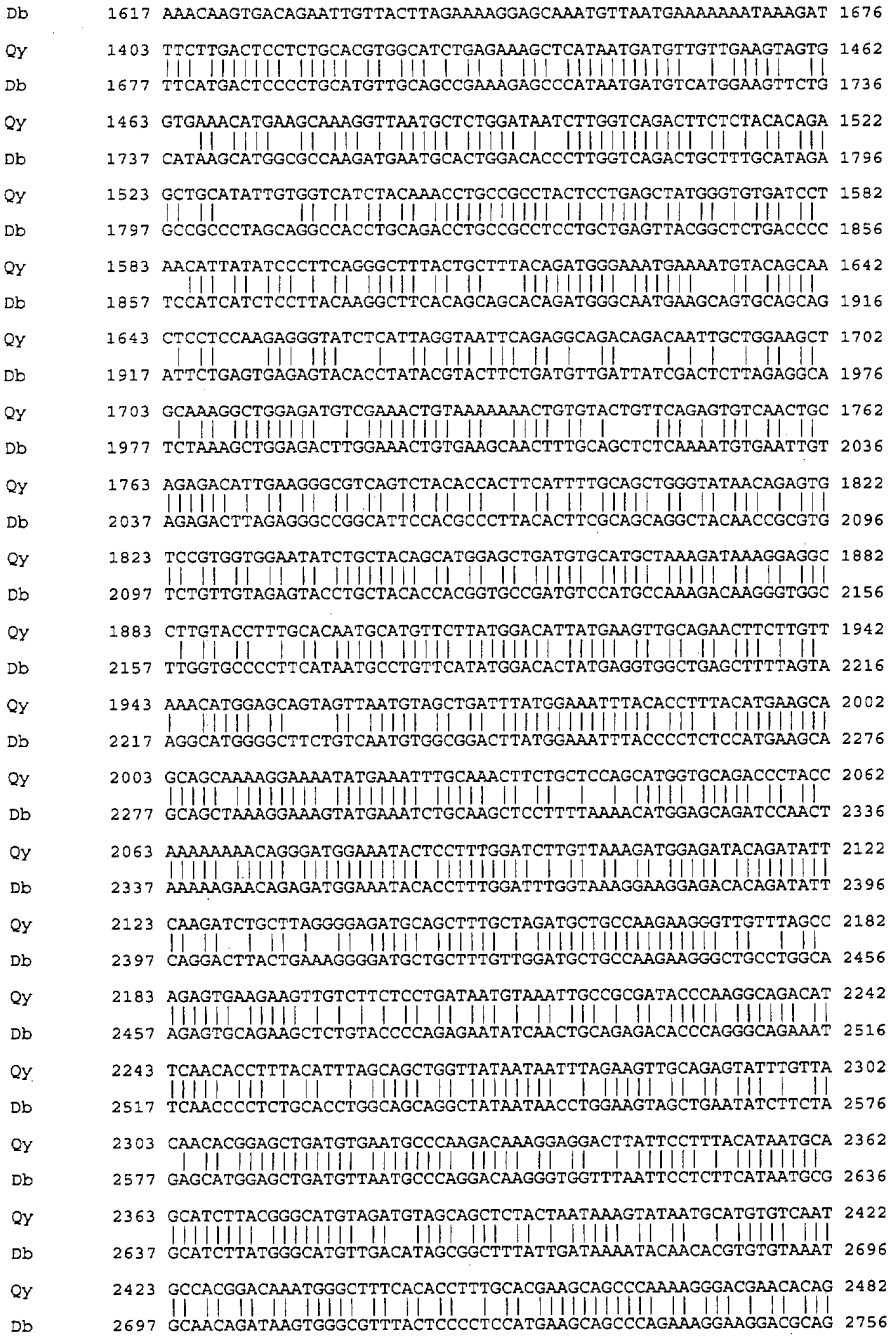
Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282

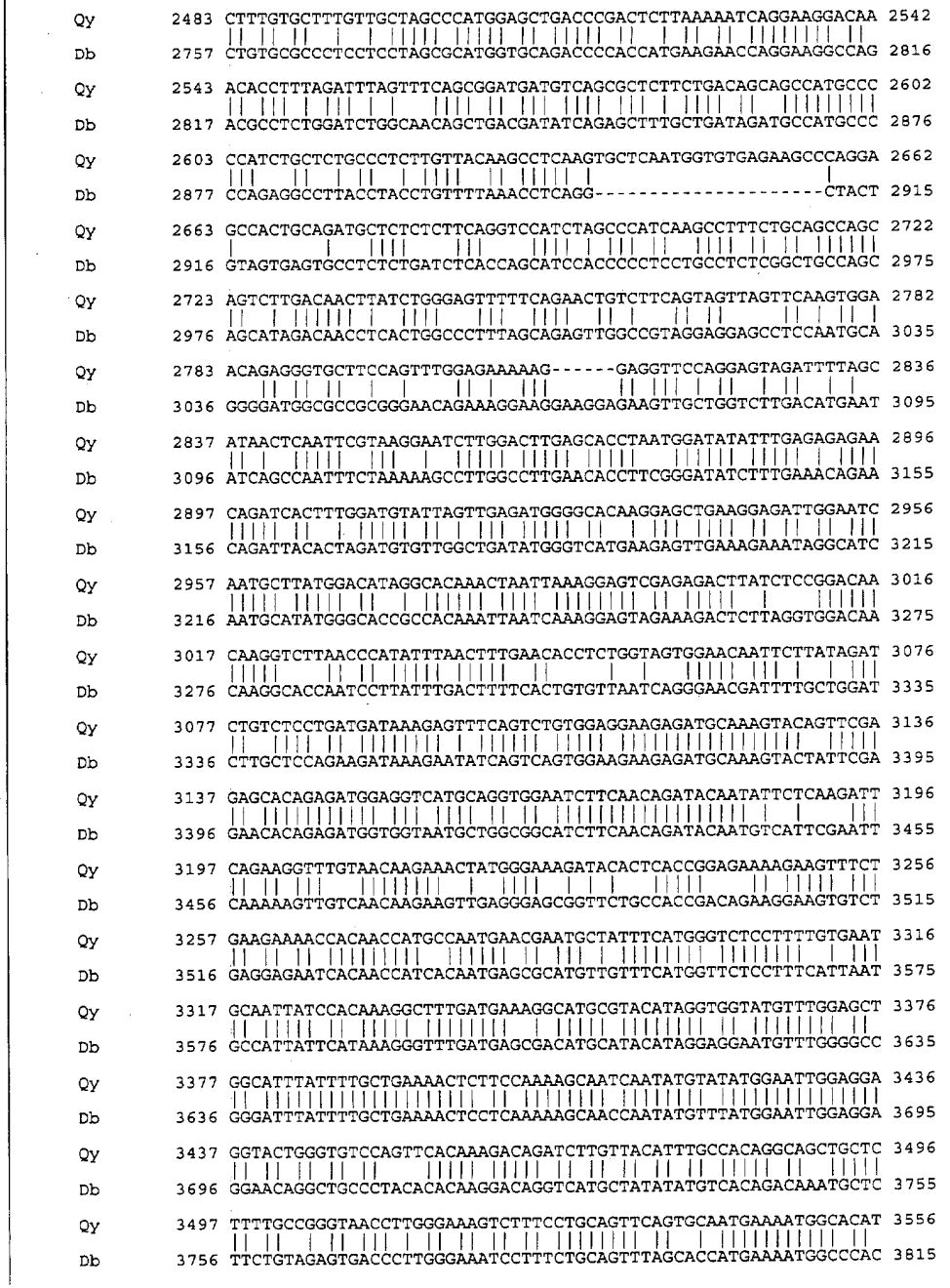
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342

Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
```





```
Qy       3557 TCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCT 3616
Db       3816 GCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCT 3875

Qy       3617 GAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATT 3676
Db       3876 GAATATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATC 3935

Qy       3677 ATGAGGCCTGAAG 3689
Db       3936 ATGAAGCCAGAAG 3948


RESULT 12
US-09-196-387-7
; Sequence 7, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4491 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2027
US-09-196-387-7

  Query Match            38.2%;  Score 1449.6;  DB 3;  Length 4491;
  Best Local Similarity  65.3%;  Pred. No. 0;
  Matches 2397;  Conservative    0;  Mismatches  889;  Indels  384;  Gaps    3;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042
Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102
Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162
Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222
Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

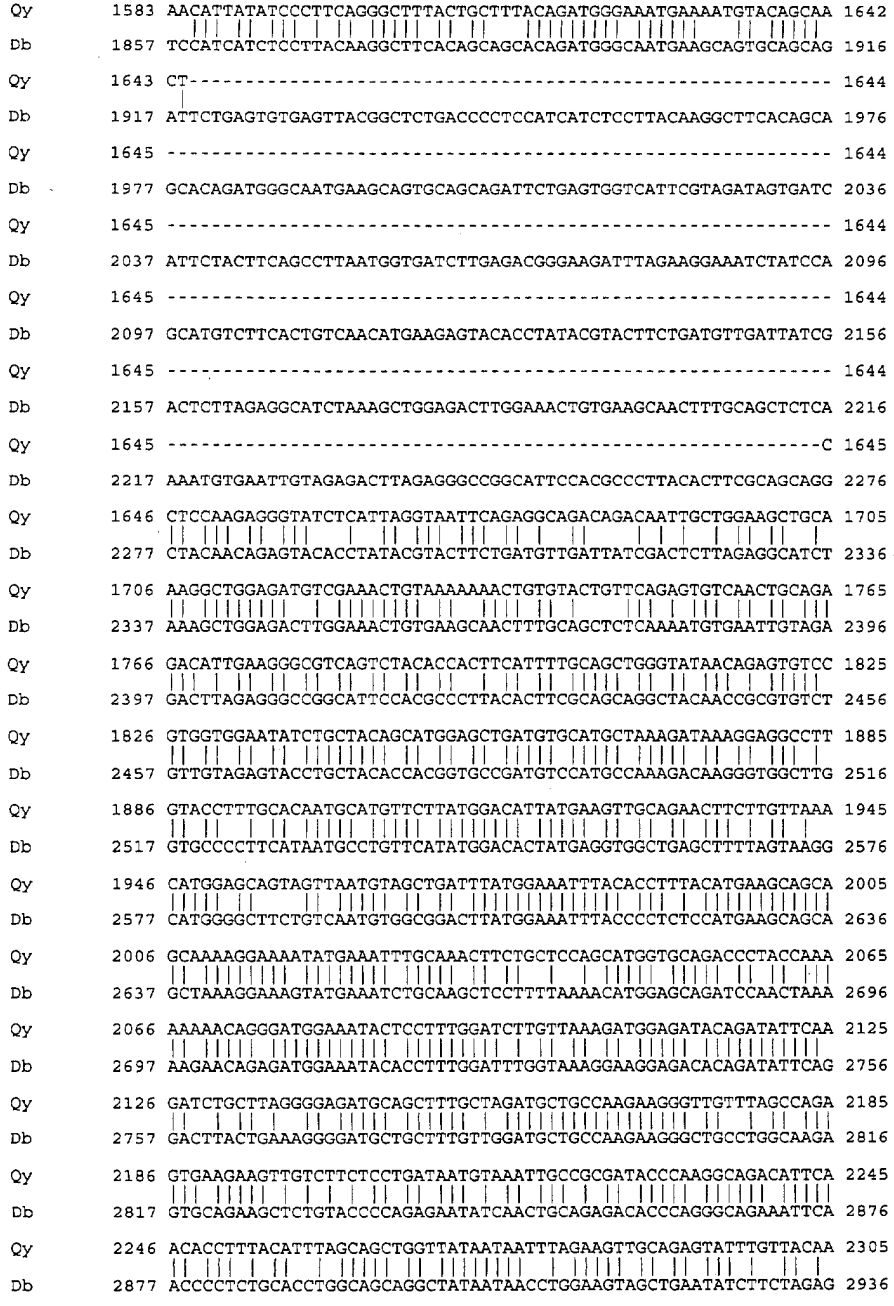
Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

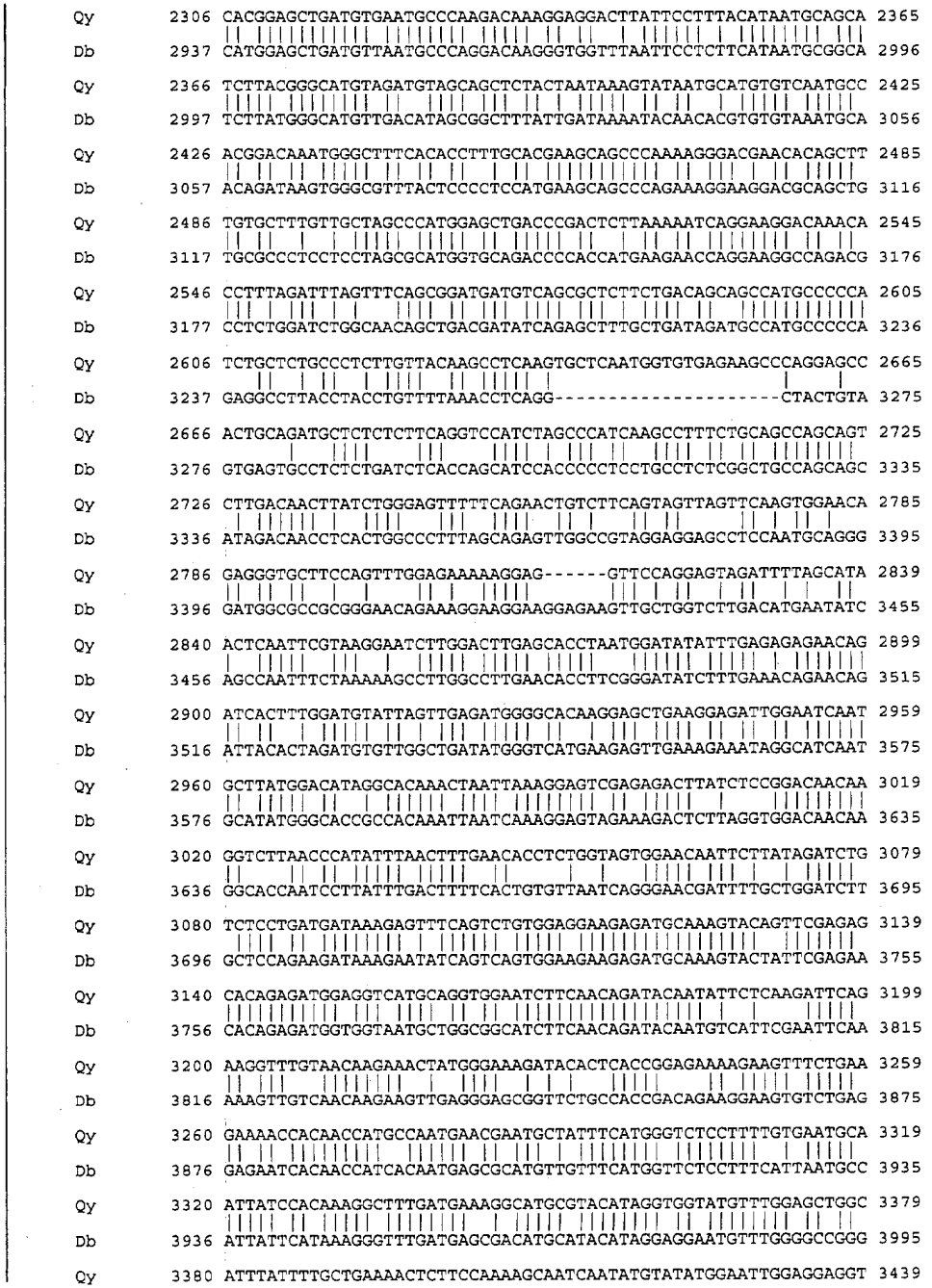
Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676

Qy       1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db       1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736

Qy       1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db       1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796

Qy       1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
Db       1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856
```





```
Db       3996 ATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGA 4055

Qy       3440 ACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTT 3499

Db       4056 ACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTC 4115

Qy       3500 TGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCT 3559

Db       4116 TGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCG 4175

Qy       3560 CCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAA 3619

Db       4176 CCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAA 4235

Qy       3620 TATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATG 3679

Db       4236 TATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATG 4295

Qy       3680 AGGCCTGAAG 3689

Db       4296 AAGCCAGAAG 4305
```

RESULT 13
US-09-841-835-7
; Sequence 7, Application US/09841835
; Patent No. 6506587
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Smith, Susan
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS
; TITLE OF INVENTION: OF USE THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841,835
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/196,387
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4491 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6..2027
US-09-841-835-7

```
  Query Match           38.2%;  Score 1449.6;  DB 4;  Length 4491;
  Best Local Similarity 65.3%;  Pred. No. 0;
  Matches 2397;  Conservative    0;  Mismatches  889;  Indels  384;  Gaps    3;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682

Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742

Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802

Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862

Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922

Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982

Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042

Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102

Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162

Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222

Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282

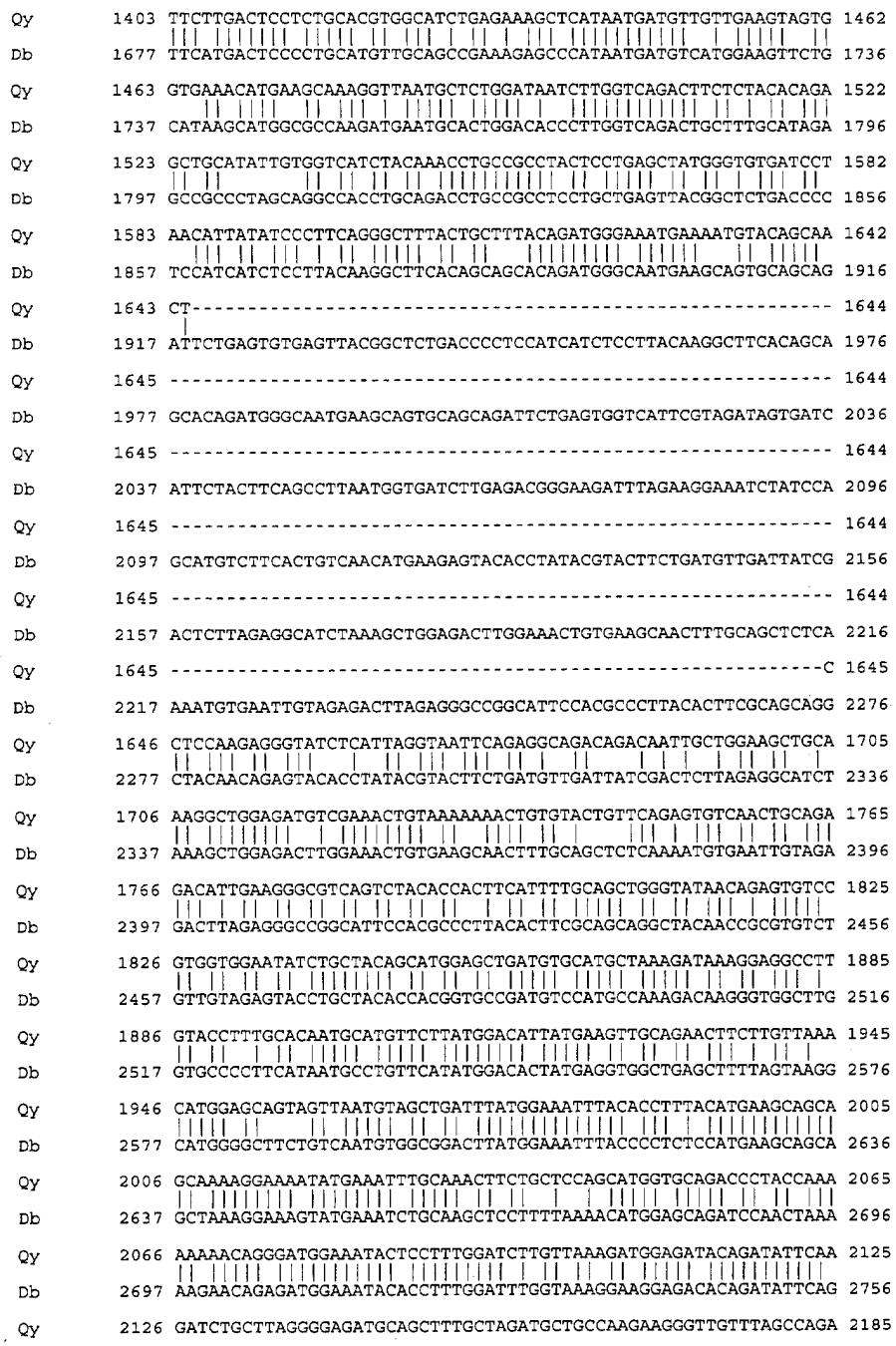
Db       1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556

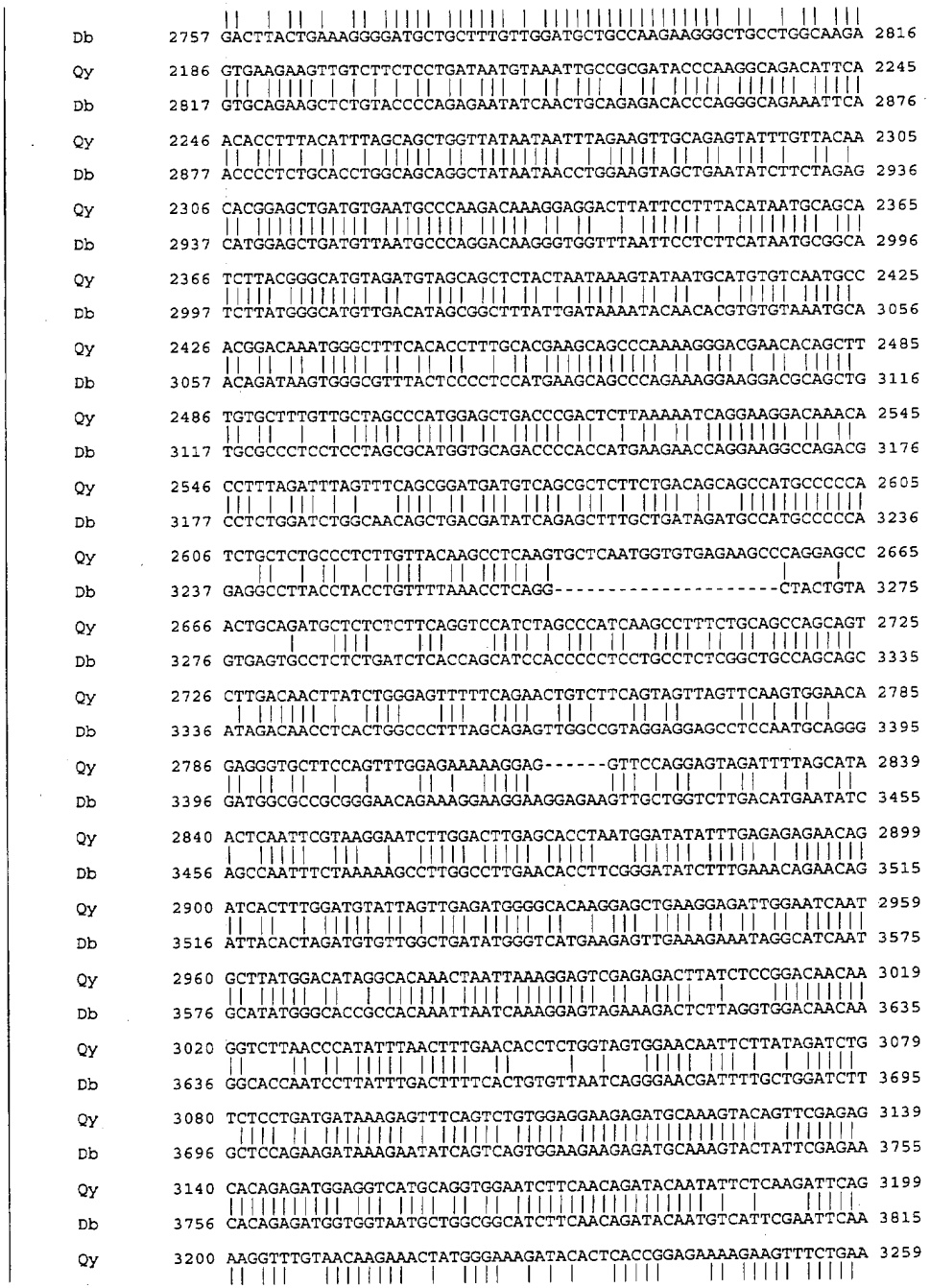
Qy       1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342

Db       1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616

Qy       1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402

Db       1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676
```





```
Db       3816 AAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTGCCACCGACAGAAGGAAGTGTCTGAG 3875

Qy       3260 GAAAACCACAACCATGCCAATGAACGAATGCTATTTCATGGGTCTCCTTTTGTGAATGCA 3319

Db       3876 GAGAATCACAACCATCACAATGAGCGCATGTTGTTTCATGGTTCTCCTTTCATTAATGCC 3935

Qy       3320 ATTATCCACAAAGGCTTTGATGAAAGGCATGCGTACATAGGTGGTATGTTTGGAGCTGGC 3379

Db       3936 ATTATTCATAAAGGGTTTGATGAGCGACATGCATACATAGGAGGAATGTTTGGGGCCGGG 3995

Qy       3380 ATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCAATATGTATATGGAATTGGAGGAGGT 3439

Db       3996 ATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCAATATGTTTATGGAATTGGAGGAGGA 4055

Qy       3440 ACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTACATTTGCCACAGGCAGCTGCTCTTT 3499

Db       4056 ACAGGCTGCCCTACACACAAGGACAGGTCATGCTATATATGTCACAGACAAATGCTCTTC 4115

Qy       3500 TGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTTCAGTGCAATGAAAATGGCACATTCT 3559

Db       4116 TGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTTTAGCACCATGAAAATGGCCCACGCG 4175

Qy       3560 CCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAGTGTAAATGGCCTAGCATTAGCTGAA 3619

Db       4176 CCTCCAGGGCACCACTCAGTCATTGGTAGACCGAGCGTCAATGGGCTGGCATATGCTGAA 4235

Qy       3620 TATGTTATTTACAGAGGAGAACAGGCTTATCCTGAGTATTTAATTACTTACCAGATTATG 3679

Db       4236 TATGTCATCTACAGAGGAGAACAGGCATACCCAGAGTATCTTATCACTTACCAGATCATG 4295

Qy       3680 AGGCCTGAAG 3689

Db       4296 AAGCCAGAAG 4305


RESULT 14
US-09-196-387-9
; Sequence 9, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4657 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2855
US-09-196-387-9

  Query Match             34.2%;  Score 1297.8;  DB 3;  Length 4657;
  Best Local Similarity   62.9%;  Pred. No. 0;
  Matches 2400;  Conservative    0;  Mismatches  907;  Indels  508;  Gaps    5;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442

Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502

Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562

Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622

Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682

Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742

Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802

Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862

Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922

Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982

Db       1197 AAAGACAAAGGTGGACTTGTGCCTCTTCATAATGCATGTTCATATGGACATTATGAAGTC 1256

Qy        983 ACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAAATGCAATGGACTTGTGGCAATTCACT 1042

Db       1257 ACAGAACTGCTACTAAAGCATGGAGCTTGTGTTAATGCCATGGATCTCTGGCAGTTTACT 1316

Qy       1043 CCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTGAAGTATGTTCTCTTCTCTTAAGTTAT 1102

Db       1317 CCACTGCACGAGGCTGCTTCCAAGAACCGTGTAGAAGTCTGCTCTTTGTTACTTAGCCAT 1376

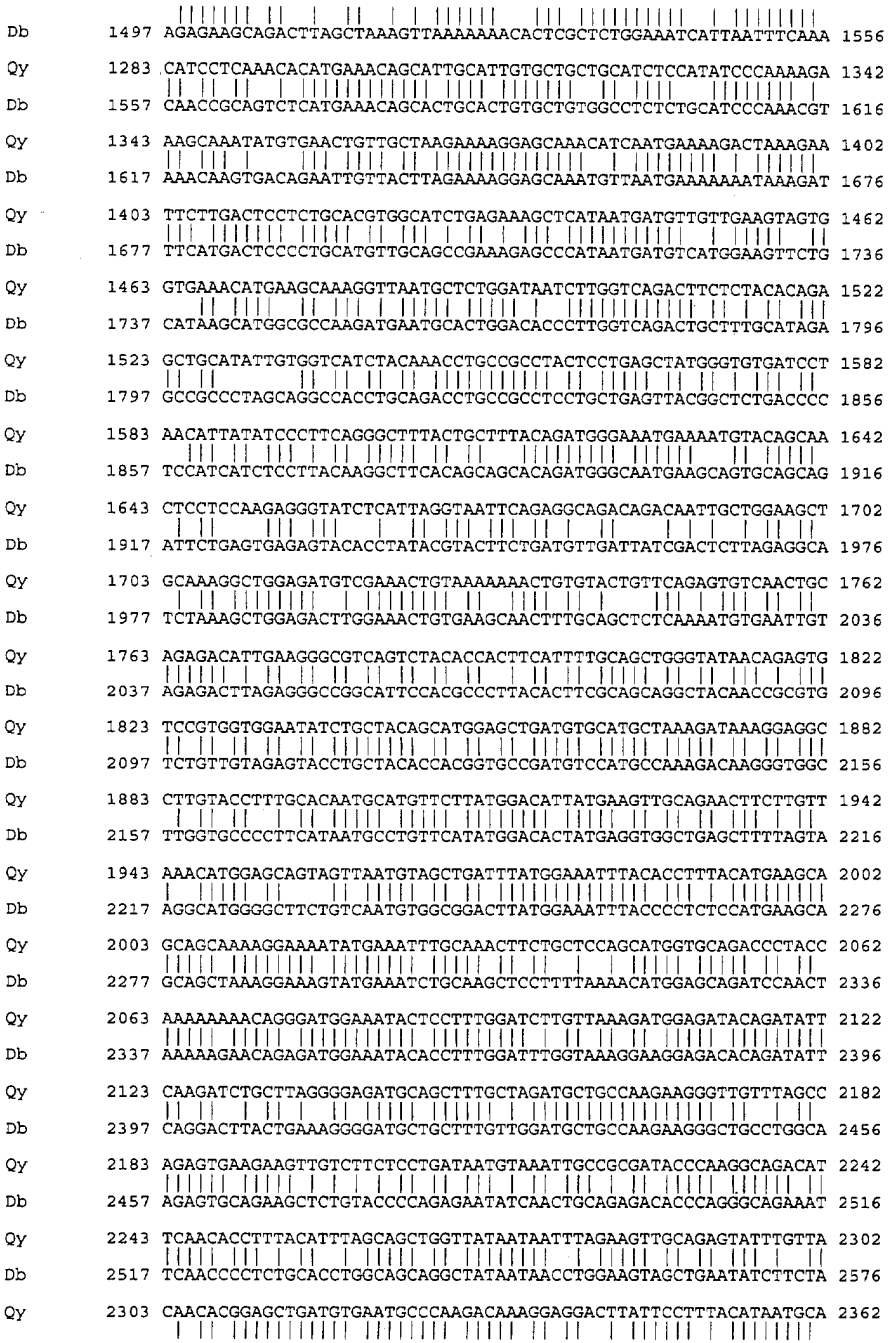
Qy       1103 GGTGCAGACCCAACACTGCTCAATTGTCACAATAAAAGTGCTATAGACTTGGCTCCCACA 1162

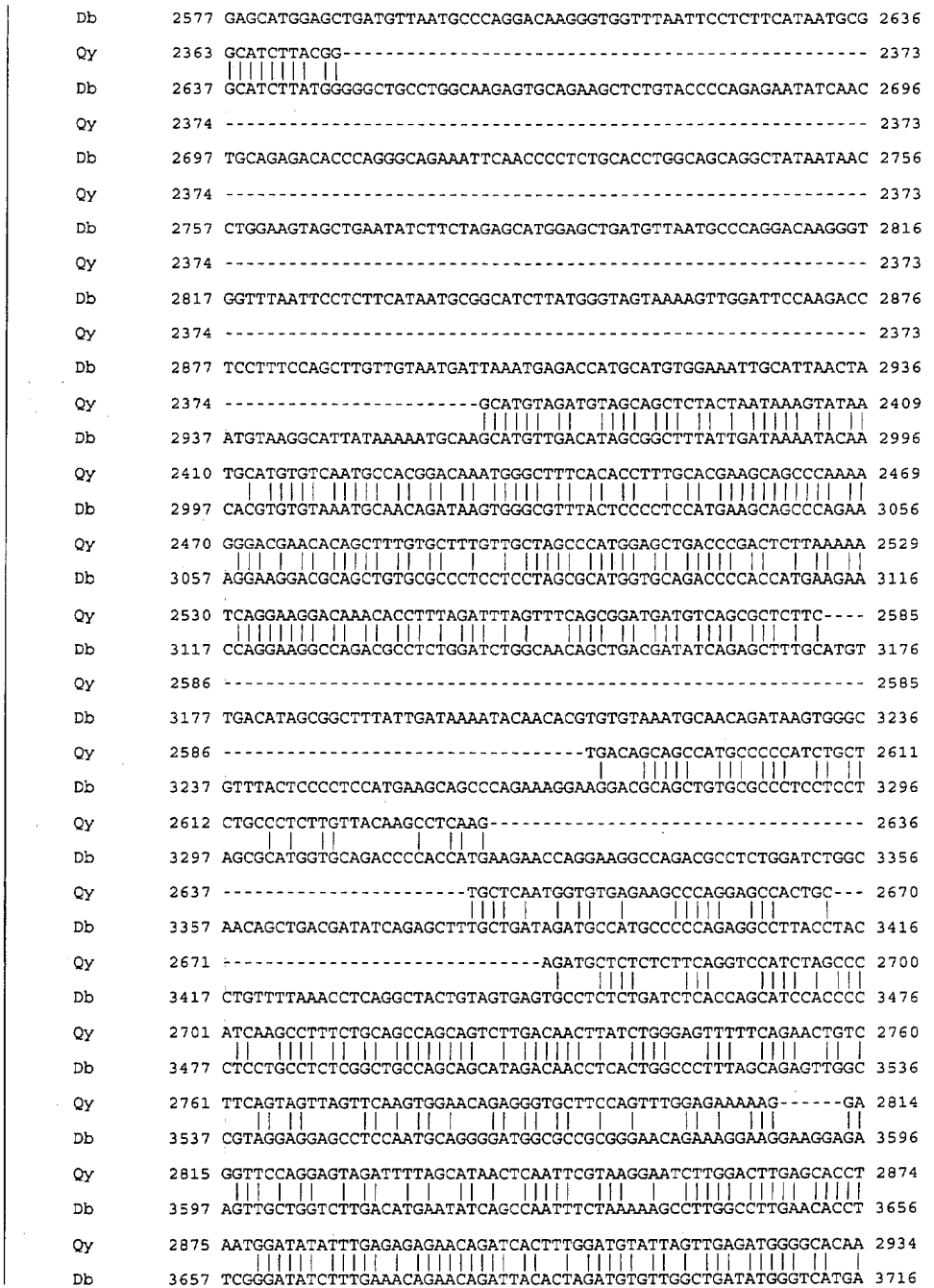
Db       1377 GGCGCTGATCCTACGTTAGTCAACTGCCATGGCAAAAGTGCTGTGGATATGGCTCCAACT 1436

Qy       1163 CCACAGTTAAAAGAAAGATTAGCATATGAATTTAAAGGCCACTCGTTGCTGCAAGCTGCA 1222

Db       1437 CCGGAGCTTAGGGAGAGATTGACTTATGAATTTAAAGGTCATTCTTTACTACAAGCAGCC 1496

Qy       1223 CGAGAAGCTGATGTTACTCGAATCAAAAAACATCTCTCTCTGGAAATGGTGAATTTCAAG 1282
```


Db 1497 AGAGAAGCAGACTTAGCTAAAGTTAAAAAAACACTCGCTCTGGAAATCATTAATTTCAAA 1556
Qy 1283 CATCCTCAAACACATGAAACAGCATTGCATTGTGCTGCTGCATCTCCATATCCCAAAAGA 1342
Db 1557 CAACCGCAGTCTCATGAAACAGCACTGCACTGTGCTGTGGCCTCTCTGCATCCCAAACGT 1616
Qy 1343 AAGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAA 1402
Db 1617 AAACAAGTGACAGAATTGTTACTTAGAAAAGGAGCAAATGTTAATGAAAAAAATAAAGAT 1676
Qy 1403 TTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTG 1462
Db 1677 TTCATGACTCCCCTGCATGTTGCAGCCGAAAGAGCCCATAATGATGTCATGGAAGTTCTG 1736
Qy 1463 GTGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGA 1522
Db 1737 CATAAGCATGGCGCCAAGATGAATGCACTGGACACCCTTGGTCAGACTGCTTTGCATAGA 1796
Qy 1523 GCTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCT 1582
Db 1797 GCCGCCCTAGCAGGCCACCTGCAGACCTGCCGCCTCCTGCTGAGTTACGGCTCTGACCCC 1856
Qy 1583 AACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAA 1642
Db 1857 TCCATCATCTCCTTACAAGGCTTCACAGCAGCACAGATGGGCAATGAAGCAGTGCAGCAG 1916
Qy 1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
Db 1917 ATTCTGAGTGAGAGTACACCTATACGTACTTCTGATGTTGATTATCGACTCTTAGAGGCA 1976
Qy 1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
Db 1977 TCTAAAGCTGGAGACTTGGAAACTGTGAAGCAACTTTGCAGCTCTCAAAATGTGAATTGT 2036
Qy 1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
Db 2037 AGAGACTTAGAGGGCCGGCATTCCACGCCCTTACACTTCGCAGCAGGCTACAACCGCGTG 2096
Qy 1823 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1882
Db 2097 TCTGTTGTAGAGTACCTGCTACACCACGGTGCCGATGTCCATGCCAAAGACAAGGGTGGC 2156
Qy 1883 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1942
Db 2157 TTGGTGCCCCTTCATAATGCCTGTTCATATGGACACTATGAGGTGGCTGAGCTTTTAGTA 2216
Qy 1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2002
Db 2217 AGGCATGGGGCTTCTGTCAATGTGGCGGACTTATGGAAATTTACCCCTCTCCATGAAGCA 2276
Qy 2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2062
Db 2277 GCAGCTAAAGGAAAGTATGAAATCTGCAAGCTCCTTTTAAAACATGGAGCAGATCCAACT 2336
Qy 2063 AAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATATT 2122
Db 2337 AAAAAGAACAGAGATGGAAATACACCTTTGGATTTGGTAAAGGAAGGAGACACAGATATT 2396
Qy 2123 CAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCC 2182
Db 2397 CAGGACTTACTGAAAGGGGATGCTGCTTTGTTGGATGCTGCCAAGAAGGGCTGCCTGGCA 2456
Qy 2183 AGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACAT 2242
Db 2457 AGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACCCAGGGCAGAAAT 2516
Qy 2243 TCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTA 2302
Db 2517 TCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCTGAATATCTTCTA 2576
Qy 2303 CAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCA 2362


Db 2577 GAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCTCTTCATAATGCG 2636
Qy 2363 GCATCTTACGG---------- 2373
Db 2637 GCATCTTATGGGGGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAAC 2696
Qy 2374 ---------- 2373
Db 2697 TGCAGAGACACCCAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAATAAC 2756
Qy 2374 ---------- 2373
Db 2757 CTGGAAGTAGCTGAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGT 2816
Qy 2374 ---------- 2373
Db 2817 GGTTTAATTCCTCTTCATAATGCGGCATCTTATGGGTAGTAAAAGTTGGATTCCAAGACC 2876
Qy 2374 ---------- 2373
Db 2877 TCCTTTCCAGCTTGTTGTAATGATTAAATGAGACCATGCATGTGGAAATTGCATTAACTA 2936
Qy 2374 ----------GCATGTAGATGTAGCAGCTCTACTAATAAAGTATAA 2409
Db 2937 ATGTAAGGCATTATAAAAATGCAAGCATGTTGACATAGCGGCTTTATTGATAAAATACAA 2996
Qy 2410 TGCATGTGTCAATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAA 2469
Db 2997 CACGTGTGTAAATGCAACAGATAAGTGGGCGTTTACTCCCCTCCATGAAGCAGCCCAGAA 3056
Qy 2470 GGGACGAACACAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAA 2529
Db 3057 AGGAAGGACGCAGCTGTGCGCCCTCCTCCTAGCGCATGGTGCAGACCCCACCATGAAGAA 3116
Qy 2530 TCAGGAAGGACAAACACCTTTAGATTTAGTTTCAGCGGATGATGTCAGCGCTCTTC---- 2585
Db 3117 CCAGGAAGGCCAGACGCCTCTGGATCTGGCAACAGCTGACGATATCAGAGCTTTGCATGT 3176
Qy 2586 ---------- 2585
Db 3177 TGACATAGCGGCTTTATTGATAAAATACAACACGTGTGTAAATGCAACAGATAAGTGGGC 3236
Qy 2586 ----------TGACAGCAGCCATGCCCCCATCTGCT 2611
Db 3237 GTTTACTCCCCTCCATGAAGCAGCCCAGAAAGGAAGGACGCAGCTGTGCGCCCTCCTCCT 3296
Qy 2612 CTGCCCTCTTGTTACAAGCCTCAAG---------- 2636
Db 3297 AGCGCATGGTGCAGACCCCACCATGAAGAACCAGGAAGGCCAGACGCCTCTGGATCTGGC 3356
Qy 2637 ----------TGCTCAATGGTGTGAGAAGCCCAGGAGCCACTGC--- 2670
Db 3357 AACAGCTGACGATATCAGAGCTTTGCTGATAGATGCCATGCCCCCAGAGGCCTTACCTAC 3416
Qy 2671 ----------AGATGCTCTCTCTTCAGGTCCATCTAGCCC 2700
Db 3417 CTGTTTTAAACCTCAGGCTACTGTAGTGAGTGCCTCTCTGATCTCACCAGCATCCACCCC 3476
Qy 2701 ATCAAGCCTTTCTGCAGCCAGCAGTCTTGACAACTTATCTGGGAGTTTTTCAGAACTGTC 2760
Db 3477 CTCCTGCCTCTCGGCTGCCAGCAGCATAGACAACCTCACTGGCCCTTTAGCAGAGTTGGC 3536
Qy 2761 TTCAGTAGTTAGTTCAAGTGGAACAGAGGGTGCTTCCAGTTTGGAGAAAAAG------GA 2814
Db 3537 CGTAGGAGGAGCCTCCAATGCAGGGGATGGCGCCGCGGGAACAGAAAGGAAGGAAGGAGA 3596
Qy 2815 GGTTCCAGGAGTAGATTTTAGCATAACTCAATTCGTAAGGAATCTTGGACTTGAGCACCT 2874
Db 3597 AGTTGCTGGTCTTGACATGAATATCAGCCAATTTCTAAAAAGCCTTGGCCTTGAACACCT 3656
Qy 2875 AATGGATATATTTGAGAGAGAACAGATCACTTTGGATGTATTAGTTGAGATGGGGCACAA 2934
Db 3657 TCGGGATATCTTTGAAACAGAACAGATTACACTAGATGTGTTGGCTGATATGGGTCATGA 3716

```
Qy       2935 GGAGCTGAAGGAGATTGGAATCAATGCTTATGGACATAGGCACAAACTAATTAAAGGAGT 2994
Db       3717 AGAGTTGAAAGAAATAGGCATCAATGCATATGGGCACCGCCACAAATTAATCAAAGGAGT 3776

Qy       2995 CGAGAGACTTATCTCCGGACAACAAGGTCTTAACCCATATTTAACTTTGAACACCTCTGG 3054
Db       3777 AGAAAGACTCTTAGGTGGACAACAAGGCACCAATCCTTATTTGACTTTTCACTGTGTTAA 3836

Qy       3055 TAGTGGAACAATTCTTATAGATCTGTCTCCTGATGATAAAGAGTTTCAGTCTGTGGAGGA 3114
Db       3837 TCAGGGAACGATTTTGCTGGATCTTGCTCCAGAAGATAAAGAATATCAGTCAGTGGAAGA 3896

Qy       3115 AGAGATGCAAAGTACAGTTCGAGAGCACAGAGATGGAGGTCATGCAGGTGGAATCTTCAA 3174
Db       3897 AGAGATGCAAAGTACTATTCGAGAACACAGAGATGGTGGTAATGCTGGCGGCATCTTCAA 3956

Qy       3175 CAGATACAATATTCTCAAGATTCAGAAGGTTTGTAACAAGAAACTATGGGAAAGATACAC 3234
Db       3957 CAGATACAATGTCATTCGAATTCAAAAAGTTGTCAACAAGAAGTTGAGGGAGCGGTTCTG 4016

Qy       3235 TCACCGGAGAAAAGAAGTTTCTGAAGAAAACCACAACCATGCCAATGAACGAATGCTATT 3294
Db       4017 CCACCGACAGAAGGAAGTGTCTGAGGAGAATCACAACCATCACAATGAGCGCATGTTGTT 4076

Qy       3295 TCATGGGTCTCCTTTTGTGAATGCAATTATCCACAAAGGCTTTGATGAAAGGCATGCGTA 3354
Db       4077 TCATGGTTCTCCTTTCATTAATGCCATTATTCATAAAGGGTTTGATGAGCGACATGCATA 4136

Qy       3355 CATAGGTGGTATGTTTGGAGCTGGCATTTATTTTGCTGAAAACTCTTCCAAAAGCAATCA 3414
Db       4137 CATAGGAGGAATGTTTGGGGCCGGGATTTATTTTGCTGAAAACTCCTCAAAAAGCAACCA 4196

Qy       3415 ATATGTATATGGAATTGGAGGAGGTACTGGGTGTCCAGTTCACAAAGACAGATCTTGTTA 3474
Db       4197 ATATGTTTATGGAATTGGAGGAGGAACAGGCTGCCCTACACACAAGGACAGGTCATGCTA 4256

Qy       3475 CATTTGCCACAGGCAGCTGCTCTTTTGCCGGGTAACCTTGGGAAAGTCTTTCCTGCAGTT 3534
Db       4257 TATATGTCACAGACAAATGCTCTTCTGTAGAGTGACCCTTGGGAAATCCTTTCTGCAGTT 4316

Qy       3535 CAGTGCAATGAAAATGGCACATTCTCCTCCAGGTCATCACTCAGTCACTGGTAGGCCCAG 3594
Db       4317 TAGCACCATGAAAATGGCCCACGCGCCTCCAGGGCACCACTCAGTCATTGGTAGACCGAG 4376

Qy       3595 TGTAAATGGCCTAGCATTAGCTGAATATGTTATTTACAGAGGAGAACAGGCTTATCCTGA 3654
Db       4377 CGTCAATGGGCTGGCATATGCTGAATATGTCATCTACAGAGGAGAACAGGCATACCCAGA 4436

Qy       3655 GTATTTAATTACTTACCAGATTATGAGGCCTGAAG 3689
Db       4437 GTATCTTATCACTTACCAGATCATGAAGCCAGAAG 4471
```

```
RESULT 15
US-09-841-835-9
; Sequence 9, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4657 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  6..2855
US-09-841-835-9

  Query Match             34.2%;  Score 1297.8;  DB 4;  Length 4657;
  Best Local Similarity   62.9%;  Pred. No. 0;
  Matches 2400;  Conservative    0;  Mismatches  907;  Indels  508;  Gaps    5;

Qy        383 TCCCTCTGGCTTTATCATTAAGGTTTTGGGCGGAAAGACGTAGTTGAATATTTGCTTCAG 442
Db        657 TCTCCCCTGCACTTCGCTGCAGGTTTTGGAAGGAAGGATGTTGTAGAACACTTACTACAG 716

Qy        443 AATGGTGCAAGTGTCCAAGCACGTGATGATGGGGGCCTTATTCCTCTTCATAATGCATGC 502
Db        717 ATGGGTGCTAATGTCCACGCTCGTGATGATGGAGGTCTCATCCCGCTTCATAATGCCTGT 776

Qy        503 TCTTTTGGTCATGCTGAAGTAGTCAATCTCCTTTTGCGACATGGTGCAGACCCCAATGCT 562
Db        777 TCTTTTGGCCATGCTGAGGTTGTGAGTCTGTTATTGTGCCAAGGAGCTGATCCAAATGCC 836

Qy        563 CGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAAGATTGATGTT 622
Db        837 AGGGATAACTGGAACTATACACCTCTGCATGAAGCTGCTATTAAAGGGAAGATCGATGTG 896

Qy        623 TGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGATGGAAGGACA 682
Db        897 TGCATTGTGCTGCTGCAGCACGGAGCTGACCCAAACATTCGGAACACTGATGGGAAATCA 956

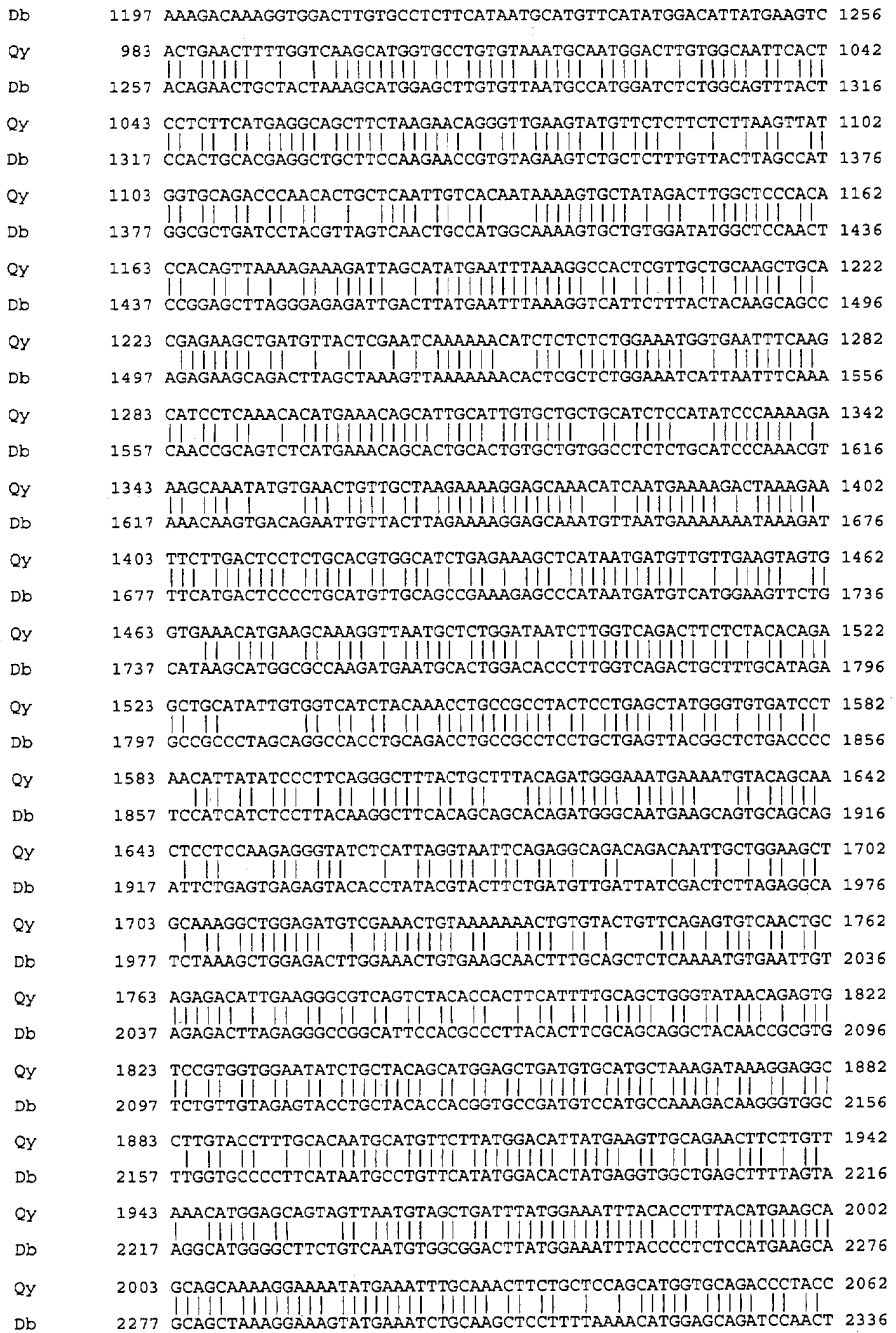
Qy        683 GCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATATAAGAAAGAT 742
Db        957 GCCCTGGACCTGGCAGATCCTTCAGCAAAAGCTGTCCTTACAGGTGAATACAAGAAAGAC 1016

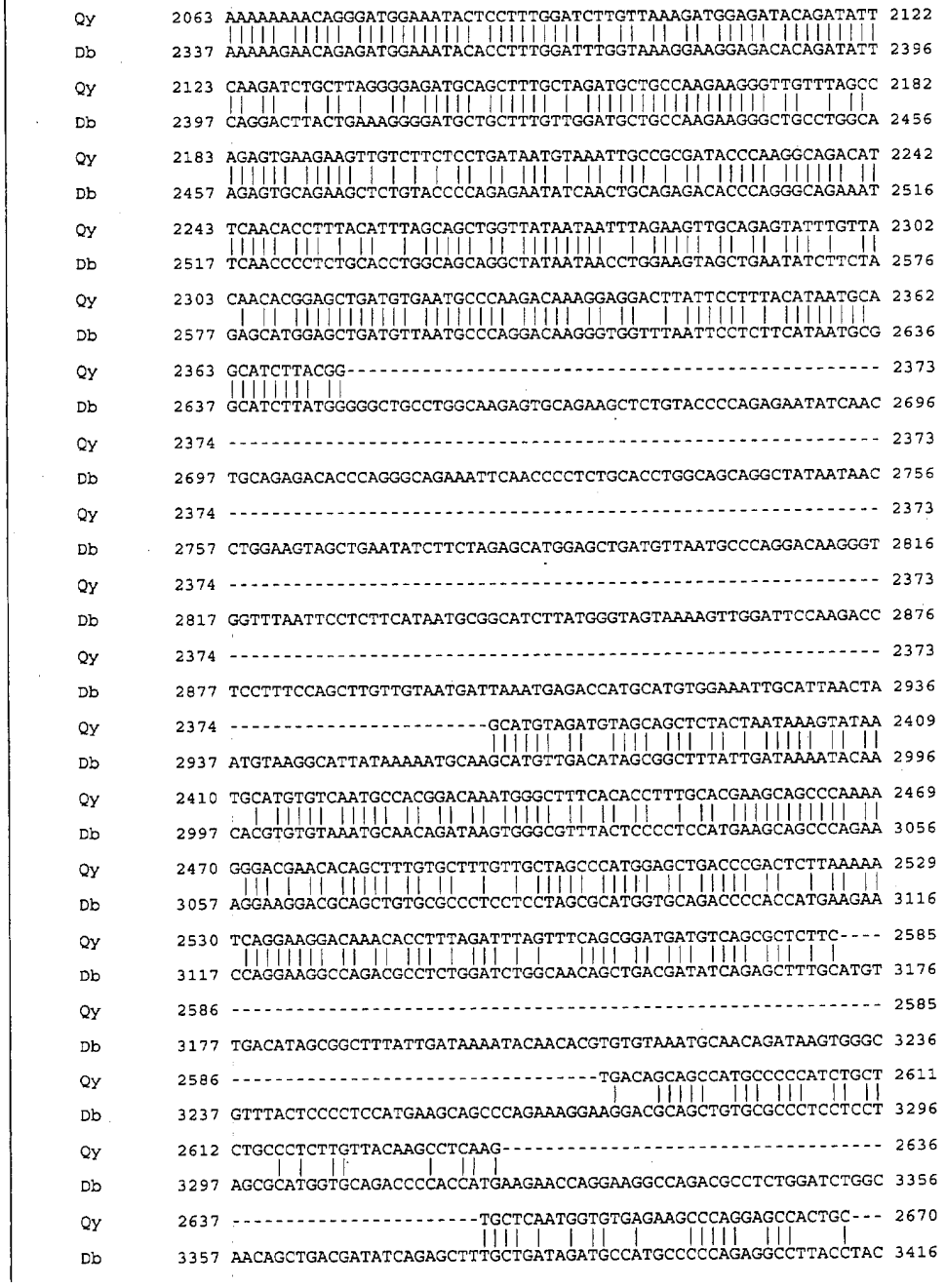
Qy        743 GAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCTACTCACACCA 802
Db       1017 GAACTCCTAGAAGCTGCTAGGAGTGGTAATGAAGAAAAACTAATGGCTTTACTGACTCCT 1076

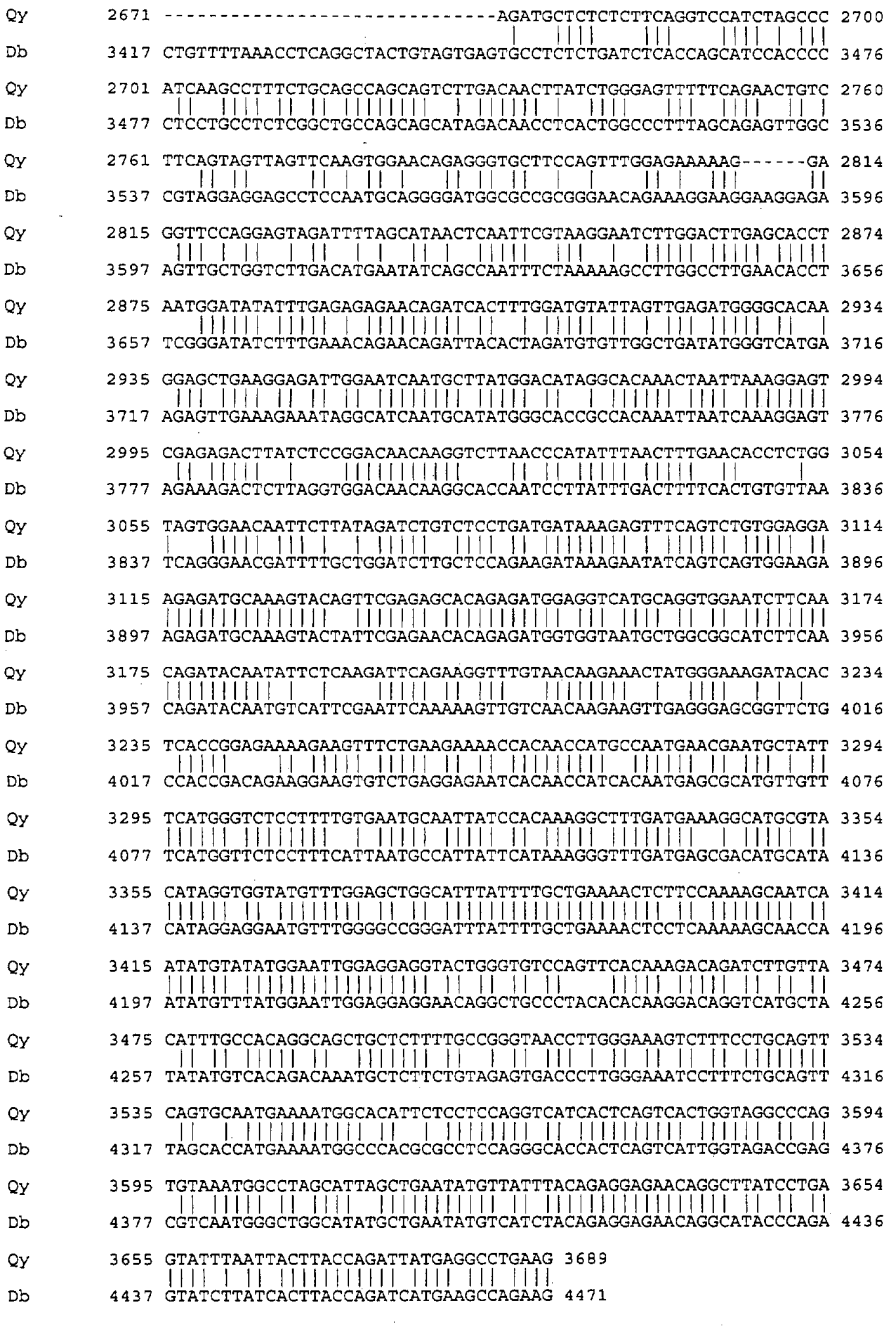
Qy        803 TTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACATTTGGCAGCA 862
Db       1077 CTAAATGTGAATTGCCATGCAAGTGATGGGCGAAAGTCGACTCCTTTACATCTAGCAGCG 1136

Qy        863 GGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGATGTCCATGCT 922
Db       1137 GGCTACAACAGAGTTCGAATAGTTCAGCTTCTTCTTCAGCATGGTGCTGATGTTCATGCA 1196

Qy        923 AAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCATTATGAAGTA 982
```







RESULT 16
US-09-972-115A-10/c
; Sequence 10, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
; APPLICANT: Geron Corporation
; APPLICANT: Gregg, Morin B.
; APPLICANT: Walter, Funk D.
; APPLICANT: Mieczyslaw, Piatyszek A.
; TITLE OF INVENTION: A Second Mammalian Telomerase
; FILE REFERENCE: 080/003C
; CURRENT APPLICATION NUMBER: US/09/972,115A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/128,577
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 60/129,123
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-972-115A-10

Query Match 13.7%; Score 518.4; DB 4; Length 520;
Best Local Similarity 99.8%; Pred. No. 4.2e-145;
Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 17
US-09-833-381-841
; Sequence 841, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

```
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 841
;   LENGTH: 1069
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1069)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-381-841

  Query Match            12.3%;  Score 467.6;  DB 4;  Length 1069;
  Best Local Similarity  98.8%;  Pred. No. 1.2e-129;
  Matches  492;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy       1643 CTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 1702
              | | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CCCTTCCTAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCT 626

Qy       1703 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 1762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 GCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGC 686

Qy       1763 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 1822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGTATAACAGAGTG 746

Qy       1823 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 1882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGC 806

Qy       1883 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 1942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTT 866

Qy       1943 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 2002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 AAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTTTACATGAAGCA 926

Qy       2003 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACC 2062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACA 986

Qy       2063 AAAAAAAACA-GGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGATACAGATAT 2121
              |||||||||| |||||||||||||||||||||||||||||||||||||| |||||||
Db        987 AAAAAAAACAGGGGATGGAAATACTCCTTTGGATCTTGTTAAAGATGGAGAT-CAGATAT 1045

Qy       2122 TCAAGATCTGCTTAGGGG 2139
              ||||||||||||||||||
Db       1046 TCAAGATCTGCTTAGGGG 1063


RESULT 18
US-09-833-381-1153
; Sequence 1153, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1153
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1153

  Query Match            12.1%;  Score 461;  DB 4;  Length 465;
  Best Local Similarity  100.0%;  Pred. No. 6.3e-128;
  Matches  461;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        553 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CCCCAATGCTCGAGATAATTGGAATTATACTCCTCTCCATGAAGCTGCAATTAAAGGAAA 64

Qy        613 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GATTGATGTTTGCATTGTGCTGTTACAGCATGGAGCTGAGCCAACCATCCGAAATACAGA 124

Qy        673 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGGAAGGACAGCATTGGATTTAGCAGATCCATCTGCCAAAGCAGTGCTTACTGGTGAATA 184

Qy        733 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 TAAGAAAGATGAACTCTTAGAAAGTGCCAGGAGTGGCAATGAAGAAAAAATGATGGCTCT 244

Qy        793 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ACTCACACCATTAAATGTCAACTGCCACGCAAGTGATGGCAGAAAGTCAACTCCATTACA 304

Qy        853 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTTGGCAGCAGGATATAACAGAGTAAAGATTGTACAGCTGTTACTGCAACATGGAGCTGA 364

Qy        913 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATGCCTGTTCTTATGGTCA 424

Qy        973 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 1013
              |||||||||||||||||||||||||||||||||||||||||
Db        425 TTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTG 465


RESULT 19
US-09-972-115A-9
; Sequence 9, Application US/09972115A
; Patent No. 6599728
; GENERAL INFORMATION:
;  APPLICANT: Geron Corporation
;  APPLICANT:  Gregg, Morin B.
;  APPLICANT:  Walter, Funk D.
;  APPLICANT:  Mieczyslaw, Piatyszek A.
;  TITLE OF INVENTION: A Second Mammalian Telomerase
;  FILE REFERENCE: 080/003C
;  CURRENT APPLICATION NUMBER: US/09/972,115A
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 60/128,577
;  PRIOR FILING DATE: 2000-04-10
;  PRIOR APPLICATION NUMBER: US 60/129,123
;  PRIOR FILING DATE: 1999-04-13
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 384
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-115A-9

  Query Match            10.0%;  Score 380.4;  DB 4;  Length 384;
  Best Local Similarity  99.7%;  Pred. No. 8.1e-104;
  Matches  381;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1682 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 1741
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          3 GACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGT 62

Qy       1742 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 1801
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTT 122

Qy       1802 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 1861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTG 182

Qy       1862 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 1921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTAT 242

Qy       1922 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 1981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAA 302

Qy       1982 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 2041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTC 362

Qy       2042 CAGCATGGTGCAGACCCTACCA 2063
              |||||||||||||||||||| |
Db        363 CAGCATGGTGCAGACCCTACAA 384
```

```
RESULT 20
US-09-196-387-12
; Sequence 12, Application US/09196387
; Patent No. 6277613
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/196,387
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/095,225
;      FILING DATE:  June 10, 1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  523 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-196-387-12
```

```
  Query Match             3.2%;  Score 121.2;  DB 3;  Length 523;
  Best Local Similarity   74.3%;  Pred. No. 5.5e-26;
  Matches  153;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy       2171 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2230
              || ||  | || |||||| ||||| | | | | || || ||| | || ||| | || |||
Db          1 GGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACC 60

Qy       2231 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2290
              || |||||| ||||||| ||| | ||  | ||||| || |||||||  | ||||| ||
Db         61 CAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCT 120

Qy       2291 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2350
              || ||| |  || | || |||||||||||| |||||||| ||||| || ||  | ||||||
Db        121 GAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCT 180

Qy       2351 TTACATAATGCAGCATCTTACGGGCA 2376
              | |||||||| |||||||| ||| |
Db        181 CTTCATAATGCGGCATCTTATGGGTA 206
```

```
RESULT 21
US-09-841-835-12
; Sequence 12, Application US/09841835
; Patent No. 6506587
;  GENERAL INFORMATION:
;    APPLICANT:  de Lange, Titia
;    APPLICANT:  Smith, Susan
;    TITLE OF INVENTION:  A PROTEIN THAT BINDS TO TRF1 AND METHODS
;    TITLE OF INVENTION:  OF USE THEREOF
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/841,835
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/196,387
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-230 CIP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  523 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
US-09-841-835-12
```

```
  Query Match             3.2%;  Score 121.2;  DB 4;  Length 523;
  Best Local Similarity   74.3%;  Pred. No. 5.5e-26;
  Matches  153;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;
```

```
Qy       2171 GGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAATTGCCGCGATACC 2230
Db          1 GGCTGCCTGGCAAGAGTGCAGAAGCTCTGTACCCCAGAGAATATCAACTGCAGAGACACC 60

Qy       2231 CAAGGCAGACATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCA 2290
Db         61 CAGGGCAGAAATTCAACCCCTCTGCACCTGGCAGCAGGCTATAATAACCTGGAAGTAGCT 120

Qy       2291 GAGTATTTGTTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCT 2350
Db        121 GAATATCTTCTAGAGCATGGAGCTGATGTTAATGCCCAGGACAAGGGTGGTTTAATTCCT 180

Qy       2351 TTACATAATGCAGCATCTTACGGGCA 2376
Db        181 CTTCATAATGCGGCATCTTATGGGTA 206


RESULT 22
US-09-220-132-30
; Sequence 30, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
;  APPLICANT:  Shyjan, Andrew W.
;  TITLE OF INVENTION:  METHODS AND COMPOSITONS FOR THE IDENTIFICATION AND ASSESSMENT
;  TITLE OF INVENTION:  OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCE
;  FILE REFERENCE:  07334-074001
;  CURRENT APPLICATION NUMBER: US/09/220,132
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  US 60/079,303
;  PRIOR FILING DATE:  1998-03-25
;  PRIOR APPLICATION NUMBER:  US 60/068,821
;  PRIOR FILING DATE:  1997-12-24
;  NUMBER OF SEQ ID NOS:  191
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 14770
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-220-132-30

  Query Match             2.6%;  Score 97.4;  DB 4;  Length 14770;
  Best Local Similarity   44.8%;  Pred. No. 8.1e-18;
  Matches  569;  Conservative    0;  Mismatches  646;  Indels   54;  Gaps    3;

Qy       1344 AGCAAATATGTGAACTGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAAT 1403
Db       1133 AGCAGGTGGTAGAAATGTTGCTTGATCGAGCTGCCCCCATTCTTTCAAAAACCAAGAATG 1192

Qy       1404 TCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGG 1463
Db       1193 GATTATCTCCATTGCACATGGCCACACAAGGGGATCATTTAAACTGCGTCCAGCTTCTCC 1252

Qy       1464 TGAAACATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAG 1523
Db       1253 TCCAGCATAATGTACCCGTGGATGATGTCACCAATGACTACCTGACTGCCCTACACGTGG 1312

Qy       1524 CTGCATATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTA 1583
Db       1313 CTGCCCACTGTGGCCATTACAAAGTTGCCAAGGTTCTCTTGGATAAGAAAGCTAACCCCA 1372

Qy       1584 ACATTATATCCCTTCAGGGCTTTACTGCTTTACAGATGGGA----------------A 1625
Db       1373 ATGCCAAAGCCCTGAATGGCTTTACCCCTCTTCATATCGCCTGCAAGAAGAATCGAATTA 1432

Qy       1626 ATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACA 1685
Db       1433 AAGTAATGGAACTCCTTCTGAAACACGGTGCATCCATCCAAGCTGTAACCGAGTCGGGCC 1492

Qy       1686 GACAATTGCTGGAAGCTGCAAAGGCTGGAGATGTCGAAACTGTAAAAAAACTGTGTACTG 1745
Db       1493 TTACCCCAATCCATGTTGCTGCCTTCATGGGGCATGTAAATATTGTATCACAACTAATGC 1552

Qy       1746 TTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAG 1805

Db       1553 ATCATGGAGCCTCACCAAACACCACCAATGTGAGAGGAGAAACAGCACTGCACATGGCAG 1612

Qy       1806 CTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATG 1865
Db       1613 CTCGCTCCGGCCAAGCTGAAGTTGTGCGGTATCTGGTACAAGACGGAGCTCAGGTAGAAG 1672

Qy       1866 CTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAG 1925
Db       1673 CTAAAGCTAAGGATGACCAAACACCACTCCACATTTCAGCCCGACTGGGGAAAGCAGACA 1732

Qy       1926 TTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTA 1985
Db       1733 TAGTACAACAGCTGTTGCAGCAAGGGGCATCTCCAAATGCAGCCACAACTTCTGGGTACA 1792

Qy       1986 CACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGC 2045
Db       1793 CCCCACTTCACCTTTCCGCCCGAGAGGGGCATGAGGATGTGGCCGCGTTCCTTTTGGATC 1852

Qy       2046 ATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCT---TGTTA 2102
Db       1853 ATGGAGCGTCTTTATCTATAACAACAAAGAAAGGATTTACTCCTCTTCATGTGGCAGCAA 1912

Qy       2103 AAGATGGAGATACAGATATTCAAGATCTGCTTAGGGGAGATGCAGCTTTGCTAGATGCTG 2162
Db       1913 AATATGGAAAGCTTGAAGTCGCCAATCTCCTGCTACAGAAAAGTGCATCTCCAGATGCTG 1972

Qy       2163 CCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTCTCCTGATAATGTAAA----- 2217
Db       1973 CTGGGAAGAGCGGGCTAACACCACTGCATGTAGCTGCACATTACGATAATCAGAAAGTGG 2032

Qy       2218 ----------------------------TTGCCGCGATACCCAAGGCAGACATTCAACAC 2249
Db       2033 CCCTTCTGCTTTTGGACCAAGGAGCCTCACCTCACGCAGCCGCAAAGAATGGTTATACGC 2092

Qy       2250 CTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACG 2309
Db       2093 CACTGCACATCGCTGCCAAAAAGAACCAGATGGACATAGCGACAACTCTGCTGGAATATG 2152

Qy       2310 GAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTT 2369
Db       2153 GTGCTGATGCCAACGCAGTTACCCGGCAAGGAATTGCTTCCGTCCATCTCGCAGCTCAGG 2212

Qy       2370 ACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGG 2429
Db       2213 AAGGGCACGTGGACATGGTGTCGCTGCTCCTCGGTAGAAATGCGAATGTGAACCTGAGCA 2272

Qy       2430 ACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTG 2489
Db       2273 ATAAGAGCGGCCTGACCCCACTCCATTTGGCTGCTCAAGAAGATCGAGTGAATGTGGCAG 2332

Qy       2490 CTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTT 2549
Db       2333 AAGTCCTCGTAAACCAAGGGGCTCATGTGGACGCCCAGACAAAGATGGGATACACACCAC 2392

Qy       2550 TAGATTTAG 2558
Db       2393 TGCATGTGG 2401


RESULT 23
US-09-031-485-35
; Sequence 35, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
```

```
;       ADDRESSEE:  Heska Corporation
;       STREET:  1825 Sharp Point Drive
;       CITY:  Fort Collins
;       STATE:  Colorado
;       COUNTRY:  USA
;       ZIP:  80525
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
;       SOFTWARE:  WordPerfect for Windows, Version 7.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/031,485
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/847,429
;       FILING DATE:  24-APR-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Verser, Carol Talkington
;       REGISTRATION NUMBER:  37,459
;       REFERENCE/DOCKET NUMBER:  HW-5
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  970/493-7272
;       TELEFAX:  970/484-9505
;   INFORMATION FOR SEQ ID NO:  35:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5235 nucleotides
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-09-031-485-35

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db        178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 237

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db        238 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 297

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db        298 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 357

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db        358 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 417

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db        418 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 477

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db        478 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 537

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db        538 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 596

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db        597 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db        657 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db        717 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 776

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db        777 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 836

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAAACAGGGATGGAAATAC 2085
Db        837 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 896

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db        897 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 956

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db        957 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db        1017 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 1061

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db        1062 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1121

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db        1122 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1181

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db        1182 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db        1242 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1301

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db        1302 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1361

Qy       2563 A 2563
Db        1362 A 1362

RESULT 24
US-09-031-485-36/c
; Sequence 36, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;     APPLICANT:  Tang, Liang
;     APPLICANT:  Blehm, E. Scot
;     TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;     TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;     TITLE OF INVENTION:  USES THEREOF
;     NUMBER OF SEQUENCES:  85
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Carol Talkington Verser, Ph.D.
;       ADDRESSEE:  Heska Corporation
;       STREET:  1825 Sharp Point Drive
;       CITY:  Fort Collins
;       STATE:  Colorado
;       COUNTRY:  USA
;       ZIP:  80525
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  Windows 95
```

```
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-031-485-36

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db       5058 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 4999

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db       4998 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 4939

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db       4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 4879

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db       4878 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 4819

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db       4818 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4759

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db       4758 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4699

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db       4698 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4640

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db       4639 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085

Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       4219 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 4175

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

Qy       2563 A 2563
              |
Db       3874 A 3874


RESULT 25
US-08-847-429A-35
; Sequence 35, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-35

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db        178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 237

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db        238 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 297

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db        298 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 357

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db        358 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 417

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db        418 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 477

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db        478 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 537

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db        538 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 596

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db        597 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db        657 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db        717 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 776

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db        777 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 836

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db        837 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 896

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db        897 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 956

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db        957 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       1017 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 1061

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       1062 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1121

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       1122 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1181

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       1182 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       1242 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1301

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       1302 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1361

Qy       2563 A 2563
Db       1362 A 1362


RESULT 26
US-08-847-429A-36/c
; Sequence 36, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-36

  Query Match             2.4%;  Score 93;  DB 1;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
```

Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

```
Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db       5058 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 4999

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db       4998 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 4939

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db       4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 4879

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db       4878 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 4819

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db       4818 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4759

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db       4758 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4699

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db       4698 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4640

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db       4639 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       4219 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 4175

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

Qy       2563 A 2563
Db       3874 A 3874
```

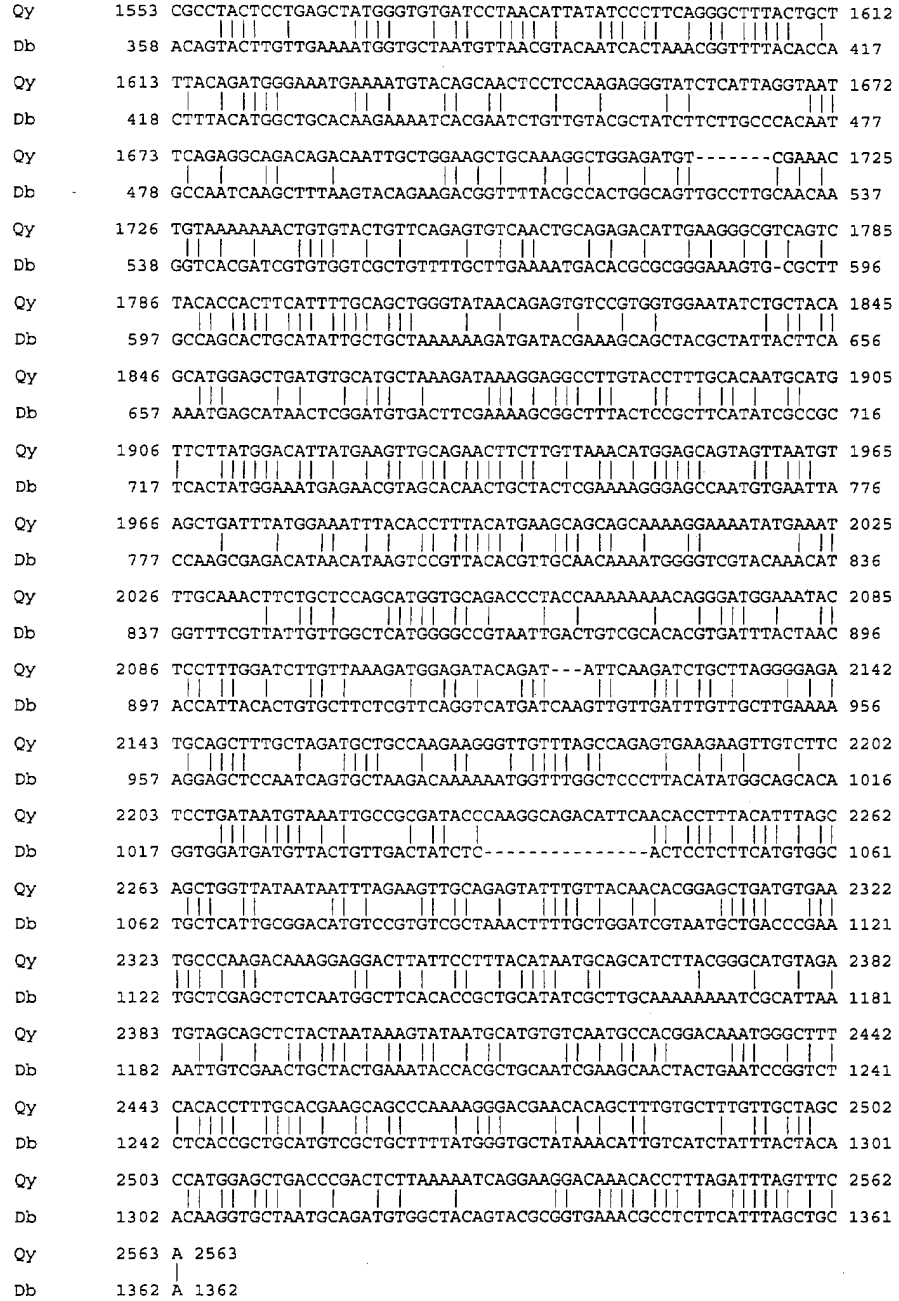
RESULT 27
US-09-065-474-35
; Sequence 35, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5235 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-065-474-35

Query Match 2.4%; Score 93; DB 3; Length 5235;
Best Local Similarity 45.0%; Pred. No. 8.1e-17;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;

```
Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db        178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 237

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db        238 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 297

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db        298 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 357
```



```
RESULT 28
US-09-065-474-36/c
; Sequence 36, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5235 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-36

  Query Match             2.4%;  Score 93;  DB 3;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;
```



```
Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db       4698 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4640

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db       4639 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       4219 GGTGGATGATGTTACTGTTGACTATCTC----------------ACTCCTCTTCATGTGGC 4175

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

Qy       2563 A 2563
Db       3874 A 3874
```

```
RESULT 29
US-09-557-034-35
; Sequence 35, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Liang
;                   Blehm, E. Scot
;        TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                            PROTEINS, NUCLEIC ACID MOLECULES, AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 171
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Carol Talkington Verser, Ph.D.
;                        Heska Corporation
;             STREET: 1825 Sharp Point Drive
;             CITY: Fort Collins
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80525
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: WordPerfect for Windows, Version 7.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,034
;             FILING DATE: 21-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/065,474
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Verser, Carol Talkington
;             REGISTRATION NUMBER: 37,459
;             REFERENCE/DOCKET NUMBER: HW-5-C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 970/493-7272
;             TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 35:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5235 nucleotides
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-557-034-35

  Query Match             2.4%;  Score 93;  DB 4;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db        178 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 237

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db        238 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 297

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db        298 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 357

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db        358 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 417

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db        418 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 477

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db        478 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 537

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db        538 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 596

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db        597 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 656

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
```

```
Db        657 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 716

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTTAAACATGGAGCAGTAGTTAATGT 1965

Db        717 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 776

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025

Db        777 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 836

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085

Db        837 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 896

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142

Db        897 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 956

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202

Db        957 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1016

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262

Db       1017 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 1061

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322

Db       1062 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1121

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382

Db       1122 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1181

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442

Db       1182 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1241

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502

Db       1242 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1301

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562

Db       1302 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1361

Qy       2563 A 2563

Db       1362 A 1362


RESULT 30
US-09-557-034-36/c
; Sequence 36, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Liang
;                   Blehm, E. Scot
;        TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                            PROTEINS, NUCLEIC ACID MOLECULES, AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 171
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Carol Talkington Verser, Ph.D.
;                        Heska Corporation
;             STREET: 1825 Sharp Point Drive
;             CITY: Fort Collins
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80525
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: WordPerfect for Windows, Version 7.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,034
;             FILING DATE: 21-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/065,474
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Verser, Carol Talkington
;             REGISTRATION NUMBER: 37,459
;             REFERENCE/DOCKET NUMBER: HW-5-C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 970/493-7272
;             TELEFAX: 970/484-9505
;    INFORMATION FOR SEQ ID NO: 36:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5235 nucleotides
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-557-034-36

  Query Match             2.4%;  Score 93;  DB 4;  Length 5235;
  Best Local Similarity   45.0%;  Pred. No. 8.1e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432

Db       5058 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 4999

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492

Db       4998 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 4939

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552

Db       4938 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 4879

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612

Db       4878 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 4819

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672

Db       4818 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4759

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725

Db       4758 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4699

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785

Db       4698 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4640

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845

Db       4639 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4580

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905

Db       4579 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4520

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTTAAACATGGAGCAGTAGTTAATGT 1965

Db       4519 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4460

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
```

```
Db       4459 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4400

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085

Db       4399 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4340

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142

Db       4339 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4280

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202

Db       4279 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4220

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262

Db       4219 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 4175

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322

Db       4174 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4115

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382

Db       4114 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4055

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442

Db       4054 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 3995

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502

Db       3994 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 3935

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562

Db       3934 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 3875

Qy       2563 A 2563

Db       3874 A 3874
```

```
RESULT 31
US-09-031-485-32
; Sequence 32, Application US/09031485
; Patent No. 5824306
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/031,485
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429
;      FILING DATE:  24-APR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  51..5285
US-09-031-485-32

  Query Match             2.4%;  Score 93;  DB 1;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432

Db        228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492

Db        288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552

Db        348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612

Db        408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672

Db        468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725

Db        528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785

Db        588 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845

Db        647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905

Db        707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965

Db        767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826

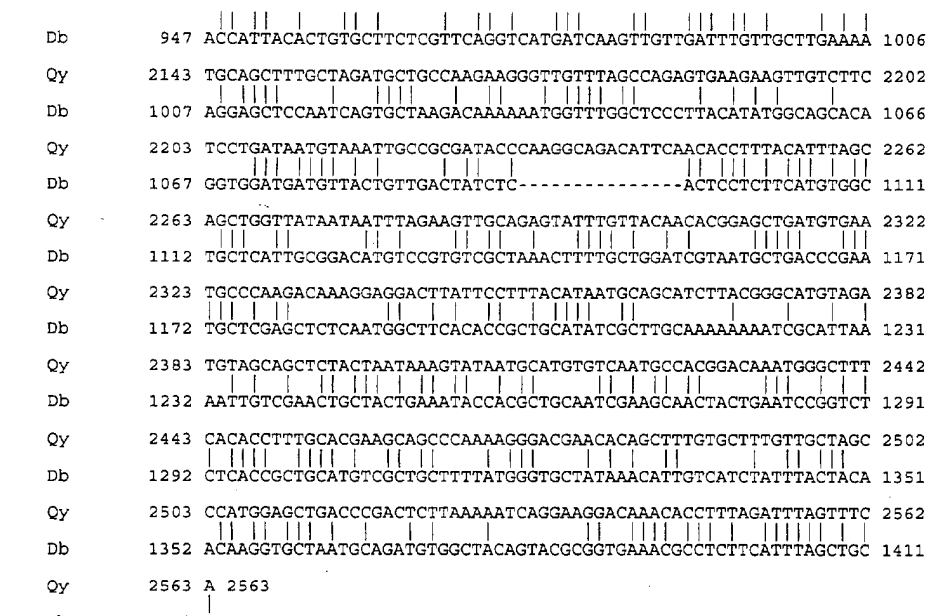
Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025

Db        827 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 886

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085

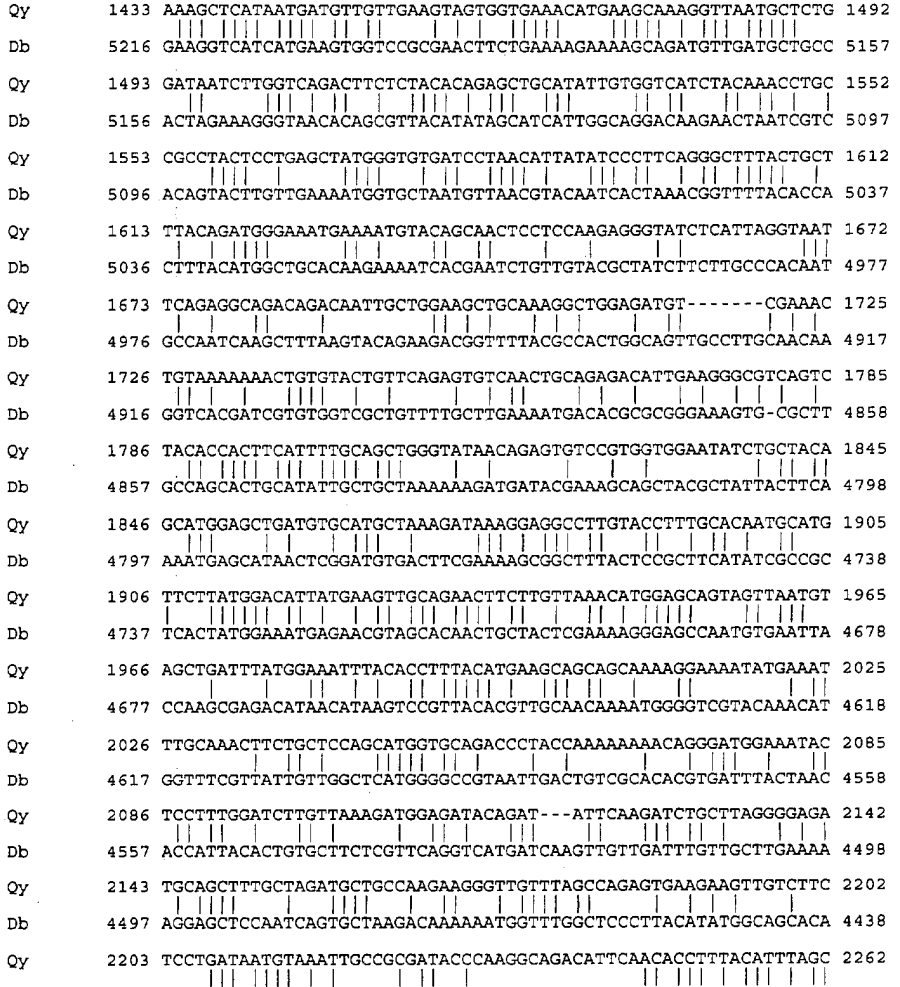
Db        887 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 946

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
```



RESULT 32
US-09-031-485-34/c
; Sequence 34, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-031-485-34

Query Match 2.4%; Score 93; DB 1; Length 5503;
Best Local Similarity 45.0%; Pred. No. 8.4e-17;
Matches 540; Conservative 0; Mismatches 635; Indels 26; Gaps 4;



```
Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2563 A 2563
              |
Db       4092 A 4092
```

```
RESULT 33
US-08-847-429A-32
; Sequence 32, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  51..5285
US-08-847-429A-32
```

```
Query Match             2.4%;  Score 93;  DB 1;  Length 5503;
Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db        228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db        288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db        348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db        408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db        468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db        528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db        588 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db        647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db        707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db        767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db        827 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 886

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db        887 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 946

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db        947 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 1006

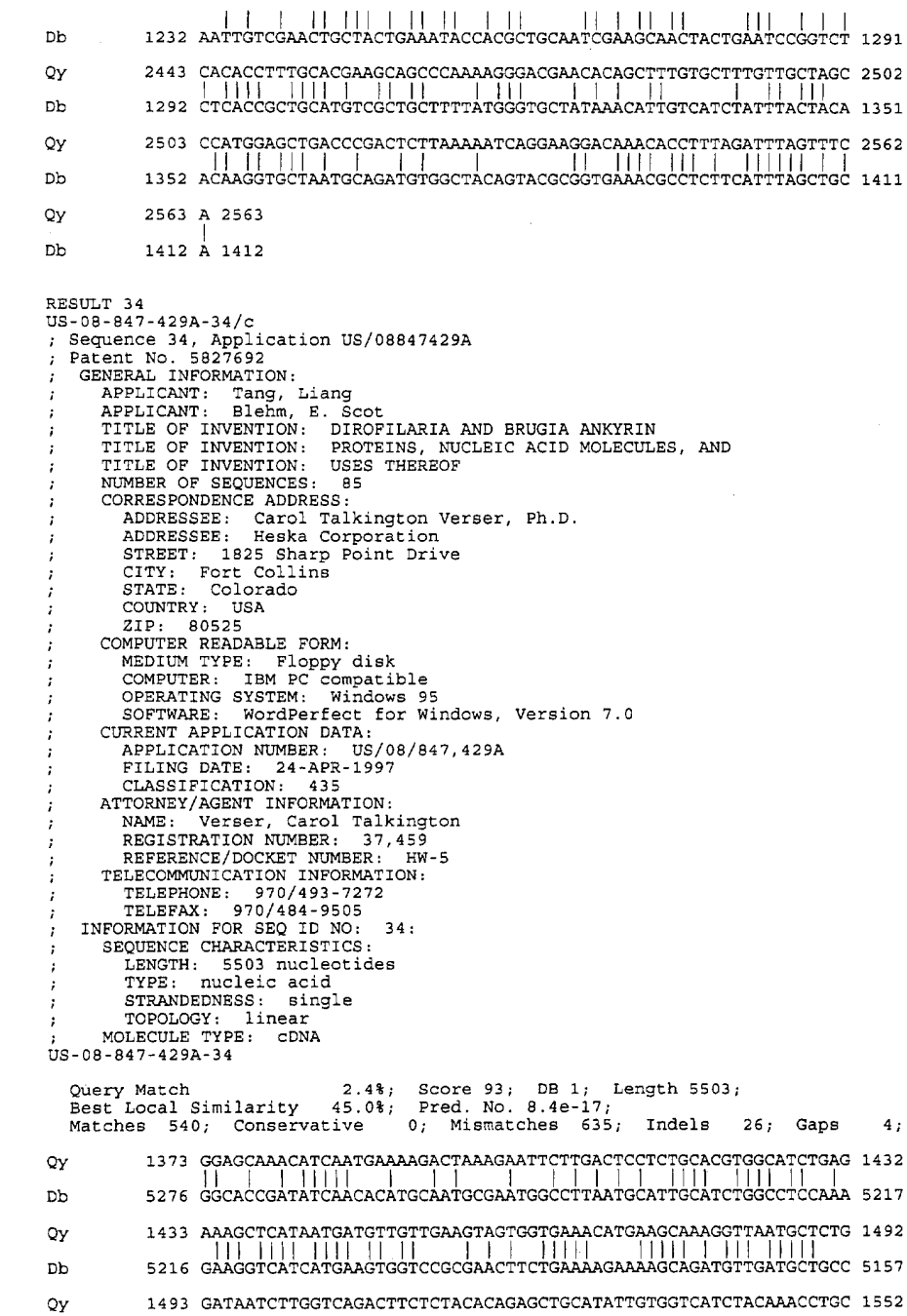
Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db        1007 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1066

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db        1067 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 1111

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db        1112 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1171

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db        1172 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1231

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
```



```
Db       1232 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1291

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502

Db       1292 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1351

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562

Db       1352 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1411

Qy       2563 A 2563

Db       1412 A 1412


RESULT 34
US-08-847-429A-34/c
; Sequence 34, Application US/08847429A
; Patent No. 5827692
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/847,429A
;      FILING DATE:  24-APR-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-847-429A-34

  Query Match             2.4%;  Score 93;  DB 1;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432

Db       5276 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 5217

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492

Db       5216 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 5157

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552

Db       5156 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 5097

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612

Db       5096 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 5037

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672

Db       5036 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4977

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT--------CGAAAC 1725

Db       4976 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4917

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785

Db       4916 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4858

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845

Db       4857 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4798

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905

Db       4797 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4738

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965

Db       4737 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4678

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025

Db       4677 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4618

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085

Db       4617 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4558

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142

Db       4557 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4498

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202

Db       4497 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4438

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262

Db       4437 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 4393

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322

Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382

Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442

Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502

Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562

Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2563 A 2563
```

```
Db       4092 A 4092


RESULT 35
US-09-065-474-32
; Sequence 32, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  51..5285
US-09-065-474-32

  Query Match             2.4%;  Score 93;  DB 3;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy      1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
             ||  |  |||||    |  | |  |  | | | | | ||||  |||| || |
Db       228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy      1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
             ||| |||| |||| || ||    | | |  |||||    ||||| | ||| |||||
Db       288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy      1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
               ||     ||| | ||  |  |||| | |||  |||       || ||   || | |
Db       348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy      1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
               ||||  |   ||||  | |||  |||| |   ||| ||  | || |||||  |
Db       408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy      1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
               |  | ||||      || |    ||  ||     |   | |           |||
Db       468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

Qy      1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
             |  |  ||      |          || | | | | | | ||      | | |
Db       528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy      1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
             ||  |  |  |||| |  |  |  | | ||  | | |  |  | | | | | | |
Db       588 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy      1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
               || |||| ||| |||| |||    |  |  |  |  |  |  |  |  | || ||
Db       647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

Qy      1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
               |||   | |  | |||  |    ||| | ||| ||  ||  | ||| |  ||
Db       707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

Qy      1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
             |  |||||| || |  | || ||| |||| ||  |  || | |||||    || |||
Db       767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826

Qy      1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
               |  |  || | | || |||||  | ||| ||  |  ||     | | ||
Db       827 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 886

Qy      2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
               |  | |  |||| || |  |  |  |  |  | | |||  |  ||
Db       887 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 946

Qy      2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
             || ||  |  || |  |  | || |  |||    ||  ||| || |  | | |
Db       947 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 1006

Qy      2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
               |||| |  |||| |  ||  | |||| || | | | | | | | |
Db       1007 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1066

Qy      2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
               ||| |||| | |    | || |                || ||| | ||| | ||
Db       1067 GGTGGATGATGTTACTGTTGACTATCTC----------------ACTCCTCTTCATGTGGC 1111

Qy      2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
               |||  ||    ||  |  || ||  |  |||| | | |  |||||   |||
Db       1112 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1171

Qy      2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
               ||| | ||     ||  | |  ||  | |||| ||          |  | | |
Db       1172 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1231

Qy      2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
               | |  |  || ||| | || ||  | ||      || | || ||    ||| | | |
Db       1232 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1291

Qy      2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
               |||||  |||| | || ||  | |||  | | | ||  | | || |||
Db       1292 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1351

Qy      2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
               || || ||| |  |  | |           || |||| ||| |  |||||  | |
Db       1352 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1411

Qy      2563 A 2563
             |
Db       1412 A 1412


RESULT 36
US-09-065-474-34/c
; Sequence 34, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
```

```
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5503 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-34

  Query Match             2.4%;  Score 93;  DB 3;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db       5276 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 5217

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db       5216 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 5157

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db       5156 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 5097

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db       5096 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 5037

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db       5036 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4977

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db       4976 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4917

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db       4916 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4858

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db       4857 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4798

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db       4797 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4738

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db       4737 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4678

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db       4677 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4618

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db       4617 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4558

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db       4557 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4498

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       4497 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4438

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       4437 GGTGGATGATGTTACTGTTGACTATCTC--------------ACTCCTCTTCATGTGGC 4393

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2563 A 2563
Db       4092 A 4092


RESULT 37
US-09-557-034-32
; Sequence 32, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
```

```
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 32:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 5503 nucleotides
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;                    (ix)   FEATURES:
;       SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-557-034-32

  Query Match             2.4%;  Score 93;  DB 4;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db        228 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 287

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db        288 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 347

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db        348 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 407

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db        408 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 467

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db        468 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 527

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT--------CGAAAC 1725
Db        528 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 587

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db        588 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 646

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db        647 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 706

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db        707 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 766

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db        767 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 826

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db        827 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 886

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db        887 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 946

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db        947 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 1006

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       1007 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 1066

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       1067 GGTGGATGATGTTACTGTTGACTATCTC----------ACTCCTCTTCATGTGGC 1111

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       1112 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 1171

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       1172 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 1231

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       1232 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 1291

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       1292 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 1351

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       1352 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 1411

Qy       2563 A 2563
Db       1412 A 1412


RESULT 38
US-09-557-034-34/c
; Sequence 34, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
```

```
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 34:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 5503 nucleotides
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-557-034-34

  Query Match             2.4%;  Score 93;  DB 4;  Length 5503;
  Best Local Similarity   45.0%;  Pred. No. 8.4e-17;
  Matches  540;  Conservative    0;  Mismatches  635;  Indels   26;  Gaps    4;

Qy       1373 GGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAG 1432
Db       5276 GGCACCGATATCAACACATGCAATGCGAATGGCCTTAATGCATTGCATCTGGCCTCCAAA 5217

Qy       1433 AAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAACATGAAGCAAAGGTTAATGCTCTG 1492
Db       5216 GAAGGTCATCATGAAGTGGTCCGCGAACTTCTGAAAAGAAAAGCAGATGTTGATGCTGCC 5157

Qy       1493 GATAATCTTGGTCAGACTTCTCTACACAGAGCTGCATATTGTGGTCATCTACAAACCTGC 1552
Db       5156 ACTAGAAAGGGTAACACAGCGTTACATATAGCATCATTGGCAGGACAAGAACTAATCGTC 5097

Qy       1553 CGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTATATCCCTTCAGGGCTTTACTGCT 1612
Db       5096 ACAGTACTTGTTGAAAATGGTGCTAATGTTAACGTACAATCACTAAACGGTTTTACACCA 5037

Qy       1613 TTACAGATGGGAAATGAAAATGTACAGCAACTCCTCCAAGAGGGTATCTCATTAGGTAAT 1672
Db       5036 CTTTACATGGCTGCACAAGAAAATCACGAATCTGTTGTACGCTATCTTCTTGCCCACAAT 4977

Qy       1673 TCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGGCTGGAGATGT-------CGAAAC 1725
Db       4976 GCCAATCAAGCTTTAAGTACAGAAGACGGTTTTACGCCACTGGCAGTTGCCTTGCAACAA 4917

Qy       1726 TGTAAAAAAACTGTGTACTGTTCAGAGTGTCAACTGCAGAGACATTGAAGGGCGTCAGTC 1785
Db       4916 GGTCACGATCGTGTGGTCGCTGTTTTGCTTGAAAATGACACGCGCGGGAAAGTG-CGCTT 4858

Qy       1786 TACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACA 1845
Db       4857 GCCAGCACTGCATATTGCTGCTAAAAAAGATGATACGAAAGCAGCTACGCTATTACTTCA 4798

Qy       1846 GCATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATG 1905
Db       4797 AAATGAGCATAACTCGGATGTGACTTCGAAAAGCGGCTTTACTCCGCTTCATATCGCCGC 4738

Qy       1906 TTCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGT 1965
Db       4737 TCACTATGGAAATGAGAACGTAGCACAACTGCTACTCGAAAAGGGAGCCAATGTGAATTA 4678

Qy       1966 AGCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAAT 2025
Db       4677 CCAAGCGAGACATAACATAAGTCCGTTACACGTTGCAACAAAATGGGGTCGTACAAACAT 4618

Qy       2026 TTGCAAACTTCTGCTCCAGCATGGTGCAGACCCTACCAAAAAAAACAGGGATGGAAATAC 2085
Db       4617 GGTTTCGTTATTGTTGGCTCATGGGGCCGTAATTGACTGTCGCACACGTGATTTACTAAC 4558

Qy       2086 TCCTTTGGATCTTGTTAAAGATGGAGATACAGAT---ATTCAAGATCTGCTTAGGGGAGA 2142
Db       4557 ACCATTACACTGTGCTTCTCGTTCAGGTCATGATCAAGTTGTTGATTTGTTGCTTGAAAA 4498

Qy       2143 TGCAGCTTTGCTAGATGCTGCCAAGAAGGGTTGTTTAGCCAGAGTGAAGAAGTTGTCTTC 2202
Db       4497 AGGAGCTCCAATCAGTGCTAAGACAAAAAATGGTTTGGCTCCCTTACATATGGCAGCACA 4438

Qy       2203 TCCTGATAATGTAAATTGCCGCGATACCCAAGGCAGACATTCAACACCTTTACATTTAGC 2262
Db       4437 GGTGGATGATGTTACTGTTGACTATCTC----------------ACTCCTCTTCATGTGGC 4393

Qy       2263 AGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTGTTACAACACGGAGCTGATGTGAA 2322
Db       4392 TGCTCATTGCGGACATGTCCGTGTCGCTAAACTTTTGCTGGATCGTAATGCTGACCCGAA 4333

Qy       2323 TGCCCAAGACAAAGGAGGACTTATTCCTTTACATAATGCAGCATCTTACGGGCATGTAGA 2382
Db       4332 TGCTCGAGCTCTCAATGGCTTCACACCGCTGCATATCGCTTGCAAAAAAAATCGCATTAA 4273

Qy       2383 TGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTCAATGCCACGGACAAATGGGCTTT 2442
Db       4272 AATTGTCGAACTGCTACTGAAATACCACGCTGCAATCGAAGCAACTACTGAATCCGGTCT 4213

Qy       2443 CACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACACAGCTTTGTGCTTTGTTGCTAGC 2502
Db       4212 CTCACCGCTGCATGTCGCTGCTTTTATGGGTGCTATAAACATTGTCATCTATTTACTACA 4153

Qy       2503 CCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGACAAACACCTTTAGATTTAGTTTC 2562
Db       4152 ACAAGGTGCTAATGCAGATGTGGCTACAGTACGCGGTGAAACGCCTCTTCATTTAGCTGC 4093

Qy       2563 A 2563
              |
Db       4092 A 4092


RESULT 39
US-09-065-474-138
; Sequence 138, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  138:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  1056 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1056
US-09-065-474-138

  Query Match             2.2%;  Score 83.4;  DB 3;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.9e-14;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2240 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2299
Db        427 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 486

Qy       2300 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAT 2359
Db        487 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 546

Qy       2360 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2419
Db        547 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 606

Qy       2420 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2479
Db        607 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 666

Qy       2480 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2539
Db        667 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 726

Qy       2540 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2572
Db        727 TTGACACCAATGCATTTATGTGCACAGGAGGAT 759


RESULT 40
US-09-065-474-140/c
; Sequence 140, Application US/09065474
; Patent No. 6063599
;  GENERAL INFORMATION:
;    APPLICANT:  Tang, Liang
;    APPLICANT:  Blehm, E. Scot
;    TITLE OF INVENTION:  DIROFILARIA AND BRUGIA ANKYRIN
;    TITLE OF INVENTION:  PROTEINS, NUCLEIC ACID MOLECULES, AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  171
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carol Talkington Verser, Ph.D.
;      ADDRESSEE:  Heska Corporation
;      STREET:  1825 Sharp Point Drive
;      CITY:  Fort Collins
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80525
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  WordPerfect for Windows, Version 7.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/065,474
;      FILING DATE:  24-APR-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verser, Carol Talkington
;      REGISTRATION NUMBER:  37,459
;      REFERENCE/DOCKET NUMBER:  HW-5-C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  970/493-7272
;      TELEFAX:  970/484-9505
;  INFORMATION FOR SEQ ID NO:  140:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1056 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-065-474-140

  Query Match             2.2%;  Score 83.4;  DB 3;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.9e-14;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2240 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2299
Db        630 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 571

Qy       2300 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTATTCCTTTACATAAT 2359
Db        570 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 511

Qy       2360 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2419
Db        510 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 451

Qy       2420 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2479
Db        450 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 391

Qy       2480 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2539
Db        390 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 331

Qy       2540 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2572
Db        330 TTGACACCAATGCATTTATGTGCACAGGAGGAT 298


RESULT 41
US-09-557-034-138
; Sequence 138, Application US/09557034
; Patent No. 6365569
;  GENERAL INFORMATION:
;       APPLICANT: Tang, Liang
;                  Blehm, E. Scot
;       TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                           PROTEINS, NUCLEIC ACID MOLECULES, AND
;                           USES THEREOF
;       NUMBER OF SEQUENCES: 171
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Carol Talkington Verser, Ph.D.
;                       Heska Corporation
;            STREET: 1825 Sharp Point Drive
;            CITY: Fort Collins
;            STATE: Colorado
;            COUNTRY: USA
;            ZIP: 80525
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: Windows 95
;            SOFTWARE: WordPerfect for Windows, Version 7.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/557,034
;            FILING DATE: 21-Apr-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/065,474
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Verser, Carol Talkington
;            REGISTRATION NUMBER: 37,459
```

```
;            REFERENCE/DOCKET NUMBER: HW-5-C1
;        TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 970/493-7272
;            TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 138:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 1056 nucleotides
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;                      (ix)   FEATURES
;        SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-557-034-138

  Query Match             2.2%;  Score 83.4;  DB 4;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.9e-14;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2240 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2299
             ||   |||||||  | |||  |||| ||   |||  |||||     | ||  |||   ||   |
Db        427 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 486

Qy       2300 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTTATTCCTTTACATAAT 2359
              || || |  ||  |||   ||   |||||    | |||     ||     | ||||||||||| |
Db        487 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 546

Qy       2360 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2419
              ||  ||      |    ||  | |||  | ||        ||   |  | |||||  |||  ||
Db        547 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 606

Qy       2420 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2479
              |||||        ||||    |  || |||||   |  ||     || |||||  ||||  |
Db        607 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 666

Qy       2480 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2539
              |  |   |||  ||  ||  | |   |||||||  |          | |       | | |||
Db        667 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 726

Qy       2540 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2572
                |||||  |  |||||  | ||  |||  |||
Db        727 TTGACACCAATGCATTTATGTGCACAGGAGGAT 759


RESULT 42
US-09-557-034-140/c
; Sequence 140, Application US/09557034
; Patent No. 6365569
;   GENERAL INFORMATION:
;        APPLICANT: Tang, Liang
;                   Blehm, E. Scot
;        TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
;                            PROTEINS, NUCLEIC ACID MOLECULES, AND
;                            USES THEREOF
;        NUMBER OF SEQUENCES: 171
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Carol Talkington Verser, Ph.D.
;                        Heska Corporation
;             STREET: 1825 Sharp Point Drive
;             CITY: Fort Collins
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80525
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: Windows 95
;             SOFTWARE: WordPerfect for Windows, Version 7.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/557,034
;             FILING DATE: 21-Apr-2000
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/065,474
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Verser, Carol Talkington
;             REGISTRATION NUMBER: 37,459
;             REFERENCE/DOCKET NUMBER: HW-5-C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 970/493-7272
;             TELEFAX: 970/484-9505
;   INFORMATION FOR SEQ ID NO: 140:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1056 nucleotides
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 140:
US-09-557-034-140

  Query Match             2.2%;  Score 83.4;  DB 4;  Length 1056;
  Best Local Similarity   53.2%;  Pred. No. 1.9e-14;
  Matches  177;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy       2240 CATTCAACACCTTTACATTTAGCAGCTGGTTATAATAATTTAGAAGTTGCAGAGTATTTG 2299
             ||   |||||||  | |||  |||| ||   |||  |||||     | ||  |||   ||   |
Db        630 CAGGTAACACCTCTGCATGTAGCGGCACATTACAATAACGACAAGGTAGCATTGTTACTT 571

Qy       2300 TTACAACACGGAGCTGATGTGAATGCCCAAGACAAAGGAGGACTTTATTCCTTTACATAAT 2359
              || || |  ||  |||   ||   |||||    | |||     ||     | ||||||||||| |
Db        570 CTAGAAAATGGTGCTTCTGCACATGCCGCTGCCAAGAATGGGTACACTCCTTTACATATT 511

Qy       2360 GCAGCATCTTACGGGCATGTAGATGTAGCAGCTCTACTAATAAAGTATAATGCATGTGTC 2419
              ||  ||      |    ||  | |||  | ||        ||   |  | |||||  |||  ||
Db        510 GCCGCGAAGAAGAATCAGATGGATATTGCTAGCACTCTCCTTCATTATAAGGCAAATGCG 451

Qy       2420 AATGCCACGGACAAATGGGCTTTCACACCTTTGCACGAAGCAGCCCAAAAGGGACGAACA 2479
              |||||        ||||    |  || |||||   |  ||     || |||||  ||||  |
Db        450 AATGCTGAAAGCAAAGCTGGCTTTACACCACTTCATCTTGCCGCCCAGGAGGGCCATCGC 391

Qy       2480 CAGCTTTGTGCTTTGTTGCTAGCCCATGGAGCTGACCCGACTCTTAAAAATCAGGAAGGA 2539
              |  |   |||  ||  ||  | |   |||||||  |          | |       | | |||
Db        390 GAAATGGCTGCGTTATTAATTGAAAATGGAGCAAAAGTTGGAGCTCAGGCAAGGAATGGC 331

Qy       2540 CAAACACCTTTAGATTTAGTTTCAGCGGATGAT 2572
                |||||  |  |||||  | ||  |||  |||
Db        330 TTGACACCAATGCATTTATGTGCACAGGAGGAT 298


RESULT 43
US-09-833-381-842/c
; Sequence 842, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 842
;   LENGTH: 466
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-842

  Query Match             2.1%;  Score 81.6;  DB 4;  Length 466;
  Best Local Similarity   95.5%;  Pred. No. 3.8e-14;
  Matches   84;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1395 CTAAAGAATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTG 1454
```

```
Db        317 CTCTAAGATTCTTGACTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTG 258

Qy       1455 AAGTAGTGGTGAAACATGAAGCAAAGGT 1482
Db        257 AAGTAGTGGTGAAACATGAAGCAAAGGT 230
```

RESULT 44
US-09-082-059-1
; Sequence 1, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devarajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identificat
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (131)..(3394)
; OTHER INFORMATION: DNA/protein segment- human kidney Ankyrin G119
US-09-082-059-1

Query Match 2.1%; Score 78.8; DB 3; Length 3454;
Best Local Similarity 43.3%; Pred. No. 1.1e-12;
Matches 541; Conservative 0; Mismatches 662; Indels 45; Gaps 2;

```
Qy        897 TGCAACATGGAGCTGATGTCCATGCTAAAGATAAAGGTGATCTGGTACCATTACACAATG 956
Db        153 TGGATAAGAAAGCTAACCCCAATGCCAAAGCCCTGAATGGCTTTACCCCTCTTCATATTG 212

Qy        957 CCTGTTCTTATGGTCATTATGAAGTAACTGAACTTTTGGTCAAGCATGGTGCCTGTGTAA 1016
Db        213 CCTGCAAGAAGAATCGAATTAAAGTAATGGAACTCCTTCTGAAACACGGTGCATCCATCC 272

Qy       1017 ATGCAATGGACTTGTGGCAATTCACTCCTCTTCATGAGGCAGCTTCTAAGAACAGGGTTG 1076
Db        273 AAGCTGTAACCGAGTCGGGCCTTACCCCAATCCATGTTGCTGCCTTCATGGGGCATGTAA 332

Qy       1077 AAGTATGTTCTCTTCTCTTAAGTTATGGTGCAGACCCAACACTGCTCAATTGTCACAATA 1136
Db        333 ATATTGTATCACAACTAATGCATCATGGAGCCTCACCAAACACCACCAATGTGAGAGGAG 392

Qy       1137 AAAGTGCTATAGACTTGGCTCCCACACCACAGTTAAAAGAAAGATTAGCATATGAATTTA 1196
Db        393 AAACAGCACTGCACATGGCAGCTCGCTCCGGCCAAGCTGAAGTTGTGCGGTATCTGGTAC 452

Qy       1197 AAGGCCACTCGTTGCTGCAAGCTGCACGAGAAGCTGATGTTACTCGAATCAAAAAACATC 1256
Db        453 AAGACGGAGCTCAGGTAGAAGCTAAAGCTAAGGATGACCAAACACCACTCCACATTTCAG 512

Qy       1257 TCTCTCTGGAAATGGTGAATTTCAAGCATCCTCAAACACATGAAACAGCATTGCATTGTG 1316
Db        513 CCCGACTGGGGAAAGCAGACATAGTACAACAGCTGTTGCAGCAAGGGGCATCTCCAAATG 572

Qy       1317 CTGCTGCATCTCCATATCCCAAAAGAAAGCAAATATGTGAAC---------- 1358
Db        573 CAGCCACAACTTCTGGGTACACCCCACTTCACCTTTCCGCCCGAGAGGGGCATGAGGATG 632

Qy       1359 ---------TGTTGCTAAGAAAAGGAGCAAACATCAATGAAAAGACTAAAGAATTCTTGA 1409
Db        633 TGGCCGCGTTCCTTTTGGATCATGGAGCGTCTTTATCTATAACAACAAAGAAAGGATTTA 692

```

```
Qy       1410 CTCCTCTGCACGTGGCATCTGAGAAAGCTCATAATGATGTTGTTGAAGTAGTGGTGAAAC 1469
Db        693 CTCCTCTTCATGTGGCAGCAAAATATGGAAAGCTTGAAGTCGCCAATCTCCTGCTACAGA 752

Qy       1470 ATGAAGCAAAGGTTAATGCTCTGGATAATCTTGGTCAGACTTCTCTACACAGAGCTGCAT 1529
Db        753 AAAGTGCATCTCCAGATGCTGCTGGGAAGAGCGGGCTAACACCACTGCATGTAGCTGCAC 812

Qy       1530 ATTGTGGTCATCTACAAACCTGCCGCCTACTCCTGAGCTATGGGTGTGATCCTAACATTA 1589
Db        813 ATTACGATAATCAGAAAGTGGCCCTTCTGCTTTTGGACCAAGGAGCCTCACCTCACGCAG 872

Qy       1590 TATCCCTTCAGGGCTTTACTGCTTTACAGATGGGAAATGAAAATGTACAGCAACTCCTCC 1649
Db        873 CCGCAAAGAATGGTTATACGCCACTGCACATCGCTGCCAAAAAGAACCAGATGGACATAG 932

Qy       1650 AAGAGGGTATCTCATTAGGTAATTCAGAGGCAGACAGACAATTGCTGGAAGCTGCAAAGG 1709
Db        933 CGACAACTCTGCTGGAATATGGTGCTGATGCCAACGCAGTTACCCGGCAAGGAATTGCTT 992

Qy       1710 CTGGAGATGTCGAAACTGTAAAA----------------AAACTGTGTACTGTTCAGA 1751
Db        993 CCGTCCATCTCGCAGCTCAGGAAGGGCACGTGGACATGGTGTCGCTGCTCCTCGGTAGAA 1052

Qy       1752 GTGTCAACTGCAGAGACATTGAAGGGCGTCAGTCTACACCACTTCATTTTGCAGCTGGGT 1811
Db       1053 ATGCGAATGTGAACCTGAGCAATAAGAGCGGCCTGACCCCACTCCATTTGGCTGCTCAAG 1112

Qy       1812 ATAACAGAGTGTCCGTGGTGGAATATCTGCTACAGCATGGAGCTGATGTGCATGCTAAAG 1871
Db       1113 AAGATCGAGTGAATGTGGCAGAAGTCCTCGTAAACCAAGGGGCTCATGTGGACGCCCAGA 1172

Qy       1872 ATAAAGGAGGCCTTGTACCTTTGCACAATGCATGTTCTTATGGACATTATGAAGTTGCAG 1931
Db       1173 CAAAGATGGGATACACACCACTGCATGTGGGCTGCCACTATGGAAATATCAAGATTGTTA 1232

Qy       1932 AACTTCTTGTTAAACATGGAGCAGTAGTTAATGTAGCTGATTTATGGAAATTTACACCTT 1991
Db       1233 ATTTCCTGCTCCAGCATTCTGCAAAAGTTAATGCCAAAACAAAGAATGGGTATACGCCAT 1292

Qy       1992 TACATGAAGCAGCAGCAAAAGGAAAATATGAAATTTGCAAACTTCTGCTCCAGCATGGTG 2051
Db       1293 TACATCAAGCAGCACAGCAGGGGCATACGCATATAATAAATGTCTTACTTCAGAACAACG 1352

Qy       2052 CAGACCCTACCAAAAAAAACAGGGATGGAAATACTCCTTTGGATCTTG 2099
Db       1353 CCTCCCCCAATGAACTCACTGTGAATGGGAATACTGCCCTTGGCATTG 1400
```

RESULT 45
US-09-291-839-3
; Sequence 3, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-068
; CURRENT APPLICATION NUMBER: US/09/291,839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-291-839-3

Query Match 2.0%; Score 74.2; DB 3; Length 2505;
Best Local Similarity 55.1%; Pred. No. 2.1e-11;
Matches 145; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

```
Qy       1787 ACACCACTTCATTTTGCAGCTGGGTATAACAGAGTGTCCGTGGTGGAATATCTGCTACAG 1846
              ||   | || ||| ||||  |       |  |     |    |  |  ||    ||| | ||
Db        406 ACTGCCCTCCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTGCAA 465

Qy       1847 CATGGAGCTGATGTGCATGCTAAAGATAAAGGAGGCCTTGTACCTTTGCACAATGCATGT 1906
              |||||||| ||||  ||  | |||||  ||       |    || |||||  | ||||
Db        466 CATGGAGCTAATGTCAATATTCAAGATGCAGTTTTTTTCACTCCATTGCATATTGCAGCG 525

Qy       1907 TCTTATGGACATTATGAAGTTGCAGAACTTCTTGTTAAACATGGAGCAGTAGTTAATGTA 1966
              |  |||||||| |  | ||  |    |||||| | |||  ||| || |  || ||||||
Db        526 TACTATGGACATGAACAGGTAACTCGCCTTCTTTTGAAATTTGGTGCTGATGTAAATGTA 585

Qy       1967 GCTGATTTATGGAAATTTACACCTTTACATGAAGCAGCAGCAAAAGGAAAATATGAAATT 2026
              || |  |      |  || |||  | ||   |||| | |||||||||    |   | |||
Db        586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATATT 645

Qy       2027 TGCAAACTTCTGCTCCAGCATGG 2049
                |||||  || |  |  | ||
Db        646 GCAAAACTCTTGATGGAAGAAGG 668
```

Search completed: March 30, 2004, 09:47:51
Job time : 303.88 secs